Open Research Online

The Open University's repository of research publications and other research outputs

A Study of the Impact of Sex and Gender upon the Perceptions and Responses of Science Teachers

Thesis

How to cite:

Spear, Margaret Ann (1985). A Study of the Impact of Sex and Gender upon the Perceptions and Responses of Science Teachers. PhD thesis The Open University.

For guidance on citations see FAQs.

© 1984 The Author

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.21954/ou.ro.0000f939

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.

oro.open.ac.uk

D 57030/85
UNRESTRICTED

A Study of the Impact of Sex and Gender upon the Perceptions and Responses of Science Teachers

by

Margaret Ann Spear, B.Sc., M.Sc., M.Ed.

A thesis presented to the Open University
in fulfilment of the requirements for
the degree of Doctor of Philosophy in
the discipline of Educational Technology

Date of submission : December 1984
Date of award : 15 February 1985

December 1984

ProQuest Number: 27777171

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 27777171

Published by ProQuest LLC (2020). Copyright of the Dissertation is held by the Author.

All Rights Reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ABSTRACT

The study investigated the role that science teachers could play in depressing the aspirations and attainment of girls studying science. Particular attention was directed to teachers' opinions, beliefs, attitudes and expectations.

It was hypothesized that teachers

(a) perceive science to be masculine,

(b) perceive differences between boys and girls which could affect science achievement,

(c) hold higher expectations for boys than for girls,

(d) assess the work, cognitive and affective attributes of boys to be superior to those of girls.

The hypotheses were tested using attitude and rating scales, and a marking exercise.

A total of 766 science teachers were involved in the study.

In the marking exercise, 339 science teachers evaluated samples of pupils' work. The work samples and their authors were rated on a number of variables. However, pupil sex was varied so that the same piece of work was presented to half of the teachers as being the work of a girl and to the remaining teachers as being the work of a boy. Work attributed to a boy was generally rated higher for scientific accuracy and understanding of principles than identical work attributed to a girl. Furthermore, boys were judged to have significantly more aptitude for science, more favourable attitudes towards science, greater interest in science, and to be more suitable for undertaking further physical science courses.

Findings from the whole study indicate that teachers do

(a) perceive science to be masculine,

(b) perceive differences between the interests, aptitudes and future roles of girls and boys,

(c) hold sex differentiated expectations,

(d) differentially value the work and personal characteristics of boys and girls.

In conclusion, the beliefs and expectations of some science teachers probably adversely affect girls' attitudes towards science and impede their academic progress.

It is recommended that teacher education must include material that enables teachers to understand and eliminate sexism in their teaching.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors, Dr. Barbara Hodgson and Dr. Janet Gale, for their guidance and support during the planning, execution and reporting of this work.

I am also grateful to Mr. R. Austin who kindly gave me advice and assistance concerning the computer analysis.

Sincere thanks are extended to the heads of schools, heads of science and education lecturers who distributed questionnaires, and to the teachers, student teachers and school children who completed questionnaires.

Throughout the project, I received much tacit support and encouragement from my husband, John Goddard. Many thanks.

The work was carried out during the tenure of an Open University Research Studentship. Besides being grateful for the financial support received, I also appreciate the many resources and facilities that were available to me through the Institute of Educational Technology.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Page

1.0 Contents

CHAPTER 1

TABLES

FIGURES

1.1 REASONS FOR THE RESEARCH

1.1.1 The problem

Over the past 15 years there has been increasing concern about the small number of females studying science and technical subjects (Ferry, 1982; Gardner, 1974; Kelly, 1976a). Many girls cease to study the physical science subjects as soon as they become optional. Then at each successive stage in the educational system the imbalance between boys and girls studying physical science subjects increases (Kelly & Weinreich-Haste, 1979). When a science subject is compulsory, girls usually choose biology. Their rejection of physical science and technical subjects means that they often follow decidedly unbalanced curricula, which inadequately prepare them to live and work in an increasingly technological society (Harding, 1982; HMI, 1979; Kelly & Weinreich-Haste, 1979). Furthermore, their restricted choice of school subjects severely limits the range of careers open to them. As Kelly (1978a) points out, girls who have a negligible scientific background are not only narrowing their own horizons, but also depriving society of a potential source of scientific 'manpower'.

Girls who do continue with science tend to under-achieve compared to boys (Kelly, 1976a, 1978b). This tendency can be put into better perspective if comparisons are made not only with the achievement of males, but also with female achievement in other fields. It is well recognised that throughout primary and much of secondary school, girls generally achieve better results than boys (Garai & Scheinfeld, 1968; Maccoby, 1967). In the science subjects an opposing trend prevails. Boys tend to achieve better results than girls.

The two trends of female under-representation and under-achievement in science generally first become apparent at the secondary school level. A convenient measure of pupils' uptake and achievements in the different school subjects is provided by national examination statistics. Murphy

Table 1.1 Entries and percentage passes in O level chemistry

	1977	1978	1979	1980	1981
Boys					
Number of entries	79856	79550	84395	86528	87196
% passing (grades A-C)	61.1	61.30	61.99	62.93	63.01
Girls					
Number of entries	40343	43851	49729	47611	55378
% passing (grades A-C	57.83	58.74	58.54	59.06	59.67

(1978) has analysed the June 1976 General Certificate of Education, Ordinary level (GCE O level) statistics for England and Wales and he found that fewer girls entered for physics and chemistry. Furthermore, they obtained poorer grades than the boys did in these two subjects. Inspection of the DES annual reports 'Statistics of Education' for a number of years confirms that fewer girls than boys enter for O level physics and chemistry examinations, and that in chemistry the girls generally obtain a poorer pass rate (Table 1.1). See Appendix 1.1 for a fuller discussion of DES examination statistics. GCE O level statistics also show that biology is the odd one out of the sciences in that more girls take the subject. But still the boys achieve higher pass rates.

The IEA* survey of science education in 19 countries (mainly western) produced similar results (Comber & Keeves, 1973). In all the countries studied, boys scored higher than girls on science tests, although there were substantial variations in the magnitude of the differences from country to country. The difference was largest in physics and least in biology. Furthermore, the gap between the sexes in attainment steadily widened throughout the years of secondary schooling.

Female under-achievement in science is particularly disturbing when one considers that girls who choose to study the physical sciences are usually a more highly selected group than the boys, with only the most academically able girls attempting these subjects (Child, 1969; Smithers & Collings, 1981). Performance in a subject is important, since success will contribute towards interest and enthusiasm in the subject, as well as increasing expectations that affect performance and attitudes towards studying. Thus if girls are constantly outscored by the boys, they can feel discouraged by their real or apparent lack of success in science. Unsatisfactory attainment, coupled with their poorer attitudes towards science (Keeves, 1973) and weaker interest in the science subjects must turn many girls from science.

* International Association for the Evaluation of Educational Achievement

1.1.2 Commonly proposed explanations

Educationalists, psychologists, sociologists and biologists have put forward various suggestions based on sex differences, in an attempt to explain why girls dislike science, and are more likely than boys to under-achieve in science. Sex differences in intellectual functioning are frequently mentioned. Extensive surveys of the literature by Oetzel (1967) and Maccoby and Jacklin (1975) have revealed several areas of intellectual functioning where sex differences appear to exist. Girls have greater verbal ability, whereas boys have greater mathematical ability and visual-spatial ability. Thus boys would seem to be better suited for the study of science. However, the work of Lewis (1964) indicates that mathematical ability and science attainment may not be inseparable. Probably spatial ability is the ability most related to science achievement (Kelly, 1976a). Certainly science specialists have better spatial ability than arts specialists (Hudson, 1966). Nevertheless, it is unlikely that sex differences in spatial ability alone can entirely account for girls' reluctance to study science.

The stereotyped personality traits of a female - dependency, conformity, lack of confidence (Broverman et al., 1972) - contrast sharply with the traits displayed by a successful male scientist - self reliance, persistence, low sociability, non-verbal bias of intelligence (Cattell & Drevdahl, 1955; Roe, 1952). Furthermore, females are more interested in people and generally rate higher than males on nurturance and affiliation items (Walberg, 1969). Thus neither girls' personality profiles nor their interests facilitate their study of the physical sciences. The biological sciences are more favoured in view of girls' preference for the study of living things (Koelsche & Newberry, 1971).

More general aspects of socialization have also been implicated in turning girls from science. Traditionally girls play with dolls and develop caring skills, whilst boys play with mechanical and electrical toys which can aid the development of spatial skills. Allocation of household jobs usually maintains this difference in emphasis between girls' activities and boys' activities (Kelly et al., 1982). Books convey a similar message. Boys are shown engaged in active pastimes and men appear in a wide range of occupations, including scientist and engineer (Tibbetts, 1975). On the other hand, girls are depicted in more passive roles, and they can look forward to wearing an apron when they grow up (Nilsen, 1976). Most boys and girls quickly learn from their family, the mass media and peer group pressures, which activities are appropriate for their sex. They also become aware that our society condones sex linked differences in motivation and achievement.

Several writers have suggested that the physical sciences are likely to have affective connotations which appeal less to females than to males. In Ormerod and Duckworth's review (1975) of pupils' attitudes to science, they suggested that girls view science as a 'male' subject, are dissuaded by the impersonality of science, and are concerned about possible undesirable applications of science. Compared to the other sciences, biology is rated higher on humanitarian and 'social benefit' scales, and is not seen as such a potential menace to human well-being (Duckworth & Entwistle, 1974). Thus females are more likely to be engaged in biology-related fields.

Lastly, there are a number of educational factors that could well contribute to girls' poor attitudes towards and achievements in science. Girls' experiences in science lessons can be adversely affected when they are outnumbered by boys. Besides being physically crowded out by the boys and sometimes deprived of their fair share of equipment and resources, the girls can also be overlooked by the teacher (Whyte, 1983b). Thus in different ways, 'minority status' can affect performance, commitment to the subject and self-estimation (Scott, 1980). Turning to teaching strategies, there is much circumstantial evidence that girls do not respond as well as boys to discovery learning situations (Babikian, 1971; Eggleston et al., 1976). An analysis of Nuffield O level results lends further support to such a view (Harding, 1973). Moreover, the work of Eggleston et al.(1976) suggests that the teaching style most effective in eliciting favourable attitudes towards science differs for girls and boys. The context of science lessons is another factor that can influence girls' perceptions of science (Kelly, 1982). Not only the approach, but the examples and analogies used, can be more appropriate for boys than for girls. If textbooks are used, they too will emphasise that science is a male domain (Samuel, 1981; Taylor, 1979; Walford, 1980, 1981). Even examinations may be sex biased, for the use of objective tests tends to produce relatively better male performance (Harding, 1979; Hoste, 1982; Murphy, 1982). Sex differences in performance can also result from the topics examined (Dwyer, 1979; Hoste, 1982) and the context of the questions set (Graf & Riddell, 1972).

1.1.3 Critique of the explanations

Over the past 10 years, explanations of girls' under-achievement in the science subjects based upon sex differences in intellectual functioning have been severely criticised, especially by feminist writers. The fact that the relationship between pupils' performance on tests designed to measure particular cognitive abilities and their attainment and interest in science are largely a matter of supposition, not proven fact (Jenkins, 1981; Saraga, 1975), has already been mentioned in the previous section. Another problem arises as a result of the assumption that performance on a task is a direct measure of a specific ability. Griffiths and Saraga (1979) point out that virtually no attention has been paid to other factors such as anxiety, motivation and expectations, which are known to affect differentially the performance of males and females. An even more fundamental issue was raised by Griffiths and Saraga when they attacked the very basis of empirical studies into sex differences in cognitive abilities. Most of these studies, either implicitly or explicitly, were conducted without reference to a theoretical framework and so were supposedly free from theoretical assumptions. However, as Griffiths and Saraga point out, since attention was directed towards sex differences rather than similarities, it was the former that were observed, reported, and upon which theories were built. If the focus of interest had been upon similarities between the sexes, then the findings and associated literature would most probably be very different.

The findings reported in the sex difference literature should generally be viewed with scepticism. Hinton (1976) accuses that many conclusions are drawn from very flimsy evidence, much of which is open to criticism. An even more dubious practice is that of supporting conclusions by providing evidence from animal studies (Griffiths & Saraga, 1979). Besides the practice of referring to questionable evidence, there are also instances of biased reporting. Fennema (1974) points to a tendency for researchers to report results that conform to a prior expectation rather than results that conform to the actual data collected. Badger (1981) cites an example of researchers reporting statistically insignificant results as important findings. The journals in which research findings are published, themselves help to exaggerate the extent of the dissimilarity between male and female cognitive abilities by favouring studies which show sex differences (Maccoby & Jacklin, 1975). Studies which find none are unlikely to be published. This selection of research for publication means that review articles on sex differences must of necessity present a distorted image of reality. However, Badger (1981) demonstrates that reviews can be even further distorted by an author elaborating and fabricating results.

Discussions of sex differences always emphasize any differences between the mean scores of males and females. The overlap in scores between the sexes is rarely referred to (Saraga, 1975). This overlap is often quite considerable. Peden (1965) reported that testing of secondary pupils revealed that 6.3% of boys and 4.2% of girls possess aptitudes suitable for engineering. These figures indicate that on the basis of intellectual aptitude alone, 40% of engineers should be female. In Britain, less than 0.5% of engineers are women (Finniston Report, 1980). Yet the Russian figure of 38% female engineers (Brown, 1968) indicates that Peden's estimate is reasonable. Thus it would appear that sex differences in cognitive abilities are not great enough to account for the differences between the educational and occupational destinies of males and females.

The origin of sex differences in intellectual functioning is the subject of much debate. It is often assumed that such differences are 'natural', i.e. biologically determined. However, the mere identification of a sex difference does not allow us to say whether it arose as a result of socialization (a learnt difference) or of biology (an innate difference) (Hannon, 1981). Reviews by Maccoby and Jacklin (1975) and Fairweather (1976) both suggest that there is very little evidence of cognitive ability differences before puberty. This would seem to imply a social rather than biological origin (Fairweather, 1976; Griffiths, 1976). However, biological explanations generally find more favour. According to Archer (1978), they are popular because they are easy to understand, and because they can be taken to imply a 'natural order' to the world and so they do not challenge the status quo. In fact, biological theories of social phenomena usually serve the political purpose of justifying the status quo (Hinton, 1976). By drawing attention away from the influence of social factors, biological explanations of sex differences inhibit attempts to correct or ameliorate existing differences in achievement between males and females. Social and educational factors are the ones that can be changed, and thus we should direct our attention to these explanations. Even if a biological basis for sex differences does exist, the manifestation of those differences is neither inevitable or inalterable. They can be reduced or exaggerated by socialization and educational practices.

The evidence in support of sex differences in personality traits does not stand up to critical examination. Sayers (1980) contends that the literature is flawed by the use of indirect evidence and the selective reporting of findings that corroborate sex role stereotypes. The fact that many of the studies have not been replicated further detracts from their dependability. A review of the psychological literature by Maccoby and Jacklin (1975) revealed that there was no sound evidence for differences in a number of personality traits. Having dismissed exaggerations, misinterpretations and even myths, they found that the remaining evidence did not justify the stereotyped beliefs about major sex differences in personality that abound in western societies.

Clear evidence of the effect of socialization is also lacking. There can be little doubt that girls and boys are treated differently, but what is less certain is the effect of this on the development of abilities and personality relevant to the study of science. The work of Alison Kelly suggests that social and cultural factors may influence girls' decisions whether or not to opt for the science subjects, but that social factors do not adequately explain girls' lower level of achievement (Kelly, 1981; Kelly & Weinreich-Haste, 1979).

Much research has been conducted into pupils' attitudes towards science. An analysis of the IEA data for 14-year-old pupils by Kelly (1978b) produced empirical support for an attitudinal explanation. It was concluded that attitudes to science do seem to have some influence on science achievement. However, since the links between attitudes and behaviour or performance are notoriously tenuous, and since attitudes are formed in response to circumstances encountered in the home, school, community and society, attitudinal explanations may not be sufficiently elemental and independent of yet other factors.

A great variety of educational factors have been proposed to account for girls' under-representation and under-achievement in science. The evidence, which is sometimes fragmentary, is widely scattered throughout the literature in time and place. Some of the evidence is contradictory or inconclusive, and much of it is unreplicated (Kelly, 1978b). Furthermore, there has been little attempt to relate individual studies to a unifying theoretical framework. Another major criticism of many educational explanations is that the causal link between sex differentiated educational experiences and sex differentiated educational outcomes has not been established. In Kelly's IEA study (1978b) very few of the teaching or school variables were consistently related to achievement in science for either sex. Relationships which did exist were generally similar for both sexes. These negative findings could have resulted from the wrong variables being investigated, or the right variables being inadequately measured. Few people would deny that schools do contribute to girls' under-achievement in science. Quantitatively and qualitatively girls learn less science in schools than boys. It could be argued that schools merely make apparent latent sex differences in pupils' cognitive aptitudes and dispositions towards science, and then maintain the resulting differentiated position and achievement of each sex in science. However, the fact that girls' representation and achievement in science declines through the years of secondary schooling would suggest that schools may well be contributing to that decline. In any case, since school factors are more accessible to direct manipulation than affective or societal factors, they deserve still more thorough and intensive study.

The first four explanations discussed above, and to a more limited extent the last explanation as well, focus upon differences between boys and girls in attempting to explain girls' under-representation and under-achievement in science. In every case the girls are found to be less suited for science studies. Their cognitive abilities are inferior, their personalities are unsuited, their upbringing and socialization is inappropriate, their attitudes are poorer, and their learning styles are unapt. Some educationalists argue that since girls are deficient in terms of the qualities required to study science, therefore it is the girls who must change and be altered. This approach to the 'Girls and Science' problem has been widely advocated over the last 15 years, but it has not proved to be startlingly successful. Rather than place the blame upon girls and expect them to change, perhaps we should instead critically examine the various agencies that help to determine pupils' attitudes towards science, their expectations of success, and their actual level of attainment. Prime amongst these agencies must be the schools.

An assumption common to the first four explanations is that the root cause of unequal science achievement lies outside of the school system. Yet both objective evidence (experimental studies) and subjective evidence (personal reports from pupils and teachers) suggest that the schools themselves can adversely affect the science education of girls. If school factors were totally unimportant then the gap between girls' and boys' uptake of science and achievements in science should remain fairly constant from school to school. Such is not the case. A survey of mixed comprehensive schools in Yorkshire found that the proportion of fourth-year girls studying physics ranged from 2% to 66%, with similar variations in chemistry (Harris, reported in Kelly, 1981). A survey of six mixed comprehensive schools in Bedfordshire, conducted during the course of this study, also revealed great variation in the uptake of science by girls and their subsequent success in O level exams. Amongst the most academically able half of the fourth-year girls in each school, the proportion studying chemistry ranged from 28% to 74%. An inspection of the O level chemistry results for the previous year showed that between 8% - 78% of the girls who entered the subject gained grade A-C passes. Such great variations in girls' uptake and achievements in science cannot be accounted for simply in terms of biological, psychological or societal explanations. Schools must be exerting their influence upon girls' aspirations and achievements.

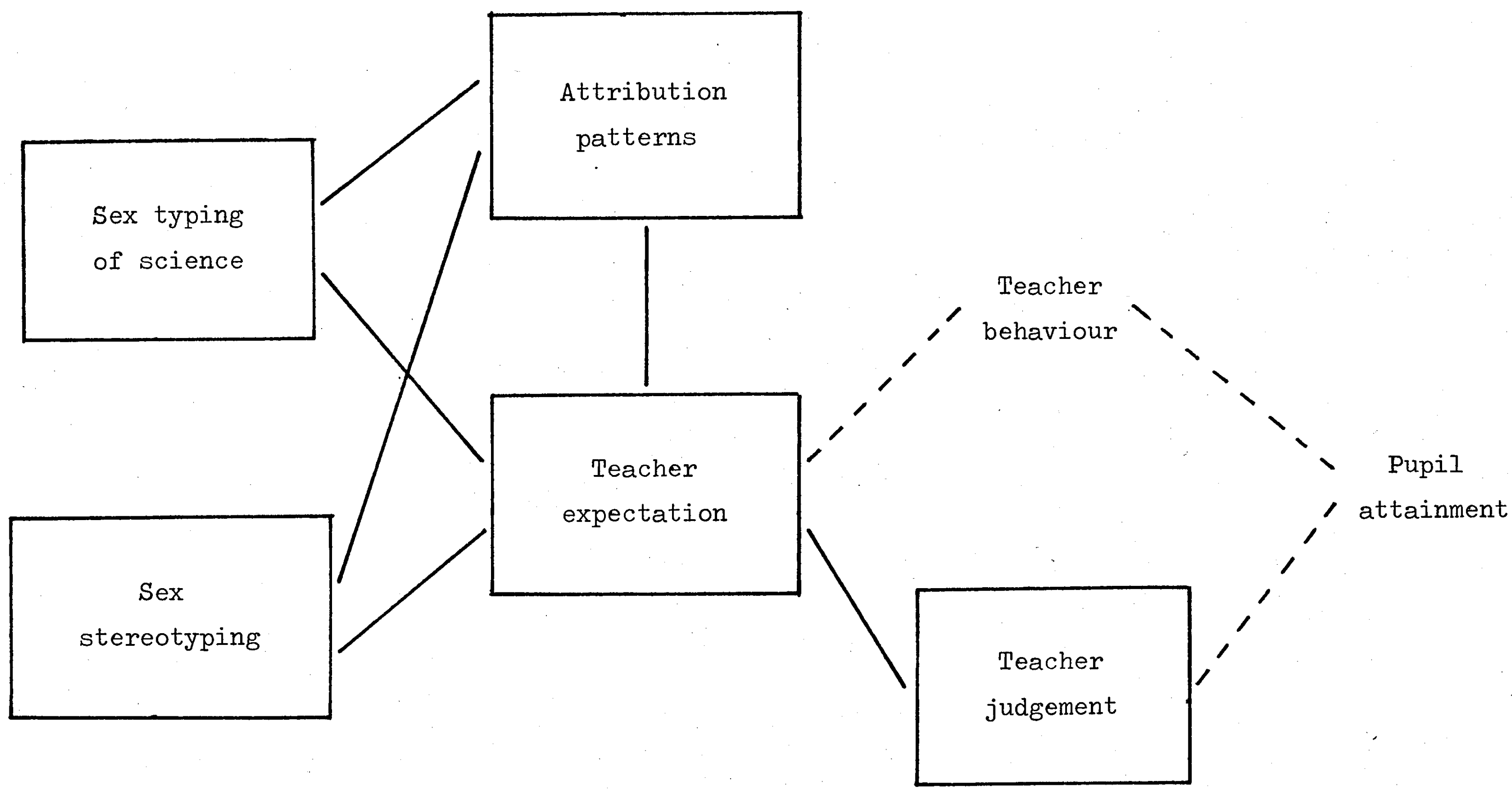
Within the school system, one crucially important factor has been largely overlooked - the teachers. Schools readily acknowledge that the teachers are of paramount importance in bringing about the education of their pupils, yet educational researchers have devoted little of their attention to teachers. Science teachers are a particularly under-researched group. Little is known, apart from anecdotal reports, about their attitudes towards girls studying science, their expectations for their female and male pupils, and their treatment of girls and boys. We need to know more about science teachers and the ways in which they respond to the sex of their pupils. In the process of filling these gaps, we may well gain fresh insights into the 'Girls and Science' problem.

1.2 AIMS AND OUTLINE OF THE RESEARCH

Most attempts to account for sex differences in science uptake and science achievement usually employ some combination of the psychological, social and educational factors mentioned above (section 1.1.2). This study focuses upon educational factors and in particular, the role that science teachers might play in discouraging girls from the physical sciences. Anecdotal reports indicate that some science teachers treat girls and boys differently. This study sets out to determine whether these claims can be substantiated and, if they can, whether this differential treatment could play a part in depressing girls' achievements and discouraging them from science.

A small scale study (Spear, 1984) has already indicated that secondary science teachers tend to display sex bias when subjectively marking samples of pupils' work. Work samples attributed to boys were generally rated higher than identical samples attributed to girls.

Figure 1.1 Model showing sex and gender effects upon science teachers' responses, perceptions and beliefs



Furthermore, 'boy' authors were generally judged to be more interested in science, and to have greater potential for physical science than 'girl' authors. The only variable for which 'girl' authors received higher ratings was neatness.

This study aims to

(a) establish whether previous findings concerning sex biased marking practices in the physical sciences are replicable,

(b) clarify which particular aspects of written work are differentially evaluated,

(c) determine whether secondary science teachers also display sex bias in their judgement of pupils' attributes relevant to success in science,

(d) investigate whether related beliefs and opinions that teachers hold about pupils are also affected by the sex of the pupil under consideration,

(e) attempt to provide an explanation for teachers' differential response to boys and to girls,

(f) identify personal and educational variables which indicate a teacher's tendency to respond differently to boys and girls.

The possibility that science teachers' sex bias arises from stereotyped beliefs about subject and pupil characteristics forms the underlying framework to this study. Figure 1.1 shows the theoretical model that was developed to guide and inform the research. If teachers believe that science is a masculine subject (i.e. they sex type* science), then they may also perceive girls as unsuited for science studies (i.e. they sex stereotype the abilities, interests, aspirations

* In the absence of any appropriate definition of the term 'sex typed' when applied to school subjects, the following understanding of the term was used as a working definition. A subject is sex typed when it is associated with one sex more than with the other, with the result that it is described and viewed as a masculine or feminine subject, i.e. it acquires a gender image.

of their pupils). These views could influence the way that teachers perceive male and female motivation towards science, and the way that teachers explain male and female success/failure in science (i.e. their attribution patterns). They could also colour the teachers' expectations regarding girls' and boys' prospects in the science subjects. Disparate teacher expectations might lead to sex biased judgements of pupil behaviour and performance, and/or differential teacher behaviour. These differential teacher behaviours could, in turn, result in differential levels of pupil attainment. Whilst it is acknowledged that the teacher expectations - pupil attainment relationship is mediated through both teacher and pupil behaviour, only teacher effects are considered in this study. The effect of sex typing of science subjects and sex stereotyping by pupils is not considered either, although both are recognised to be important determinants of a pupil's choice of science subjects. Their effect upon a pupil's level of attainment in science is more debatable. Finally, not all of the topics and links indicated in Figure 1.1 are examined. Teacher behaviour and pupil attainment, together with the links leading to them, are omitted from the present study. However, in addition to the remaining topics and relationships shown in Figure 1.1, the effect of certain teacher variables upon measures of sex typing, sex stereotyping, teacher expectation and judgement are investigated.

Besides focusing upon the role that science teachers' sex biased expectations and judgements could play in depressing the aspirations and attainment of girls in science, the study also investigates the usefulness of the model outlined above to account for those biases. In particular it needs to be established whether:

1. 'Science teachers perceive science to be masculine'

Objective data describing science teachers' views of the subjects that they teach is lacking. This study surveys the extent to which science teachers associate gender overtones with the three major science subjects as they are taught in secondary schools. It determines which aspects of school subjects contribute to gender image, inquires into how these images are maintained, and compares different ways of measuring gender images.

2. 'Science teachers perceive differences between boys and girls which could affect science achievement'

If teachers believe that the attitudes, aptitudes and roles of boys and girls are different, then they might hold different expectancies for each sex. Therefore, this study examines whether teachers do believe that the preferences, aspirations, achievements and roles of girls are different from those of boys.

3. 'Science teachers attribute boys' and girls' success/failure in science to different causes'

Some knowledge of the way that teachers use ability and effort attributions to explain boys' and girls' success/failure at science could be helpful in explaining certain aspects of teachers' expectations and behaviour, since the reasons that teachers use to explain a pupil's success will determine whether success is expected again, and the reasons used to account for failure will influence the way that teachers respond to the failure.

4. 'Pupil's sex influences teacher's expectations'

Evidence that teachers hold different expectations for boys and for girls is very sparse. Further confirmation is sought.

5. 'Pupil's sex influences teacher's judgements'

Several studies have shown that the attributes, efforts and achievements of groups perceived to be inferior are devalued. By manipulating pupil sex it should be possible to determine whether science teachers favour boys over girls in the following areas:

(a) assessment of affective factors,

(b) assessment of cognitive factors,

(c) evaluation of work.

The above topics not only guided the direction of the research reported in this thesis, but also provided structure to the thesis itself. Two of the topics, teacher expectation and teacher judgement, were effectively combined at an early stage in the research, which reduced the overall number of topics to four. Most of the chapters in this thesis are organised around these four topics.

Data were collected from teachers using questionnaires. This method enabled a large number of teachers to be contacted over a short period of time. Many of the scales which appeared on the questionnaires were developed from ideas and opinions expressed in a number of initial in-depth interviews.

Although this study primarily explored

(a) secondary science teachers' views on the masculine image of the science subjects,

(b) aspects of the sex stereotyped views held by science teachers,

(c) the extent to which science teachers overvalue the work and attributes of boys compared to those of girls,

additional data on related topics were also collected from teachers of all subjects and all levels. These supplementary data can help to place the science teachers study in context by showing

(a) how a range of secondary school subjects are viewed on a variety of characteristics, including masculinity,

(b) the extent to which teachers of other subjects distinguish between the preferences of boys and girls.

As information about teachers' attitudes towards girls and boys and their views on sex roles in general is very sparse, it is hoped that this study will provide many useful insights and explanations.

1.3 SCOPE OF THE RESEARCH

The factors that contribute to girls' under-representation and under-achievement in science are probably numerous and intricately interrelated. All of the factors mentioned in section 1.1.2 are probably important, and more besides. However, it is impossible for a single study to look at all possible causes at once. Instead it is probably most profitable to concentrate on one discrete factor, study it in detail, and attempt to assess whether it could contribute to the 'Girls and Science' problem. Although focusing upon a single factor undoubtedly oversimplifies and probably distorts our perception of the problem and its complex mesh of causal factors, nevertheless such an approach should help to advance our knowledge and understanding of the problem.

The study focuses upon science teachers (i.e. teachers of physics, chemistry, biology, integrated/general science). Of the different educational factors discussed in section 1.1.2, science teachers are arguably the single most important influence affecting girls' under-representation and under-achievement in science. To a large extent, science teachers can determine their pupils' perceptions of science, their reactions to science, and even their progress in science (see also sections 2.4 and 2.5.2). Of particular relevance to girls' under-representation in science, both survey studies and anecdotal reports indicate that science teachers can influence girls' choice of science options by the advice that they offer (Bottomley & Ormerod, 1982; DES, 1980; Reid et al., 1974), and the attitudes that they convey to girls contemplating the study of science (see section 3.1.2).

Although this study is primarily concerned with science teachers, teachers of other subjects are not totally excluded. They also convey ideas to pupils about the suitability of science for girls. The inclusion of some non-science teachers in the research not only enables the strength of their beliefs to be compared with those of science teachers, but also allows science teachers' views to be placed in a wider context.

The science teachers studied in this research all taught in secondary schools. The majority of the teachers taught in coeducational comprehensive schools, but the sample also included some teachers from boys' schools and girls' schools, and some teachers from secondary modern and grammar schools. It was inevitable that the sample would come from secondary schools since science teachers are concentrated at the secondary level. However, there were also other reasons for wishing to study teachers at the secondary level. Many girls first encounter science, particularly science as a formal discipline, when they start secondary school. Moreover, it is during secondary schooling that girls' interest in science deteriorates and the deterioration is more marked than for boys' interest in learning about science (Kelly et al., 1984). But most seriously, it is during secondary schooling that many girls choose to drop most of their science subjects; and once science studies have been discontinued, it is unlikely that they will be restarted again at a later date.

The work samples used to investigate teacher expectation and teacher judgement in this study were produced by lower secondary school pupils. At this level teachers can exert considerable influence over their pupils. The pupils have only recently transferred to a new school and so they are well motivated and generally disposed towards learning. As pupils grow older, they are less responsive to teacher influence (Good, 1980). Thus the choice of work samples from lower secondary school pupils allowed teacher effects to be investigated at a critical stage in school science education. It was also hoped that the use of lower secondary school pupils' work would minimise any tendencies on the part of teachers to respond to pupil sex on the basis of perceived societal restraints (i.e. unavailability of suitable science related jobs, negative attitudes from parents and friends), rather than intellectual restraints.

In this study, the investigations into the image of science concentrate upon the image of the three common science subjects as they are taught in secondary schools up to CSE/O level standard. The image of science in employment and research fields is not investigated.

Throughout the research, attention has been directed to the three major science subjects - physics, chemistry and biology. This decision was guided by the following considerations. The majority of secondary science teachers are trained in one of these three disciplines. The majority of secondary science teachers teach one of these three separate subjects. These are the three most popular science subjects at CSE, O level and A level. Other pure science subjects tend to be merely combinations or offshoots of these three disciplines.

In this study, a distinction is often made between the physical science subjects, i.e. physics and chemistry, and biology. It has already been noted that although girls are reluctant to study the physical sciences, they are much more willing to study biology (see Appendix 1.1). Also girls show more interest in biological topics (Bottomley & Ormerod, 1981; Smail & Kelly,in press). Ormerod & Duckworth (1975) mention several other differences between the biological and the physical sciences. Biology is regarded as an easier subject than physics or chemistry. It also has a different image. Whereas the physical sciences are seen as being impersonal, masculine and relatively unconcerned with humanitarian and social issues, the biological sciences project a more feminine image and appear to be more beneficial and less harmful to human well-being. These divergent characteristics suggest that the biological and physical science subjects can profitably be considered separately. Because girls' representation and achievement is generally worse in the physical science subjects than in biology (see Appendix 1.1), researchers have tended to focus upon girls' standing in the physical sciences. The descriptive and discursive sections of this thesis follow that custom. Wherever the term 'science' (discipline unspecified) appears, it generally refers to the two commonest physical science subjects, i.e. physics and chemistry.

The physical science and biological science subjects are usually treated separately in those sections of the thesis that specifically refer to the experimental work conducted. The inclusion of biology as well as the physical science subjects in the questionnaires widened their relevance to science teachers, and produced more comprehensive data. Furthermore, by comparing and contrasting the responses given by science teachers when referring to physical science subjects with those made when referring to biology, a better understanding of the differences and similarities between physical science and biological science subjects is possible.

CHAPTER 2

REVIEW OF THE LITERATURE

Page

2.0 Contents

CHAPTER 2

TABLES

FIGURES

2.1 THE MASCULINE IMAGE OF SCIENCE

Traditionally science has been a male preserve. This situation arose as a consequence of historical forces. In past centuries, mainly for social reasons, those engaged in scientific research and those employed in fields devoted to the application of science tended to be almost exclusively male. Science, like any other human activity, is inescapably influenced by the social order within which it operates. Thus science took on the ambience, the attitudes and the approaches of its generators and later practitioners (Mendelsohn, 1976). Most scientists, engineers and technologists were and still are men. This has resulted in science being viewed as a masculine subject.

2.1.1 Evidence of science's masculine image

There is considerable evidence, from a variety of sources, to support the contention that science has a masculine image. Most of this evidence arises from work with school pupils and university students. Adults' opinions regarding the masculine image of science have rarely been recorded systematically, but presumably the replies of school-children and students can be taken to reflect the opinions and attitudes commonly held by society.

One of the few studies involving adults was an investigation into the perceived suitability of different subjects for men and women. It is reported that 53% of a sample of American teachers (primary and secondary) stated that secondary school science is "more appropriate for a man" to teach than a women (Simpson, 1974). Such a finding provides direct evidence concerning the gender connotations of science. Using a similar technique, Cowan (1971) found that both boys and girls overwhelmingly considered science to be a subject "suitable for boys". In an international survey, Kelly (1976b) found that in several of the countries sampled, a majority of the boys considered physics to be "more suitable for boys" than for girls. Girls were less likely to subscribe to this view. According to Kelly et al. (1981), such is also the case in England. They report that boys are more likely than girls to agree with items such as "learning science is more important for boys than for girls", and to express doubt whether "girls are just as good as boys at science". In contrast, Hutt (1979) found that a higher proportion of girls than boys held the view that physics and chemistry are "better done by boys" than by girls or equally well by both. This discrepancy could be due to changes in girls' attitudes and views over the period between the two investigations, or it could have arisen as a result of Kelly's sample being involved in a large scale longitudinal study. Participation in this study will probably have enhanced the self-esteem of the pupils, and it could be that such effects have been particularly pronounced for the girls.

The semantic differential provides another useful measure of the gender associations of a subject. Feldman (1974) asked undergraduates at five American universities to rate 45 academic disciplines on a seven-point scale running from 'masculine' to 'feminine'. The physical science subjects were viewed as masculine, but the biological sciences were awarded neutral scores. Weinreich-Haste (1979) used the same method with university students in England, but embedded the 'masculine-feminine' scale within six other bipolar scales, including a 'science-arts' scale. Again physics was rated a masculine subject. Moreover, it was seen that the 'masculine' and the 'science' disciplines formed clusters and that these clusters tended to overlap.

Semantic differential studies with schoolchildren have revealed that they also perceive the physical science subjects as masculine. Working in the United States, Vockell and Lobonc (1981) found that girls attending coeducational schools view the physical sciences as more masculine than do girls at all-girl schools. In England, Weinreich-Haste (1981) reports that boys give a more masculine rating to the science subjects than do girls. Furthermore, correlations between ratings on a number of different scales suggest that science is not only viewed as masculine, but that it is also linked with other attributes that are stereotypically associated with masculinity, e.g. hardness, complexity, facts, thought, concern with objects as opposed to people (Kelly & Weinreich-Haste, 1979; Weinreich-Haste, 1981).

Ormerod (1975) has arranged school subjects into a 'gender spectrum' using the 'Brunel Subject Preference Grid'. This grid uses a paired comparison method to obtain a ranking of the popularity of a range of school subjects. The gender of a subject is determined from the sex which shows the greater preference for the subject. This method produces a 'gender spectrum' which agrees closely with Weinreich-Haste's semantic differential results. However, the position of physics and chemistry are reversed, with Weinreich-Haste showing physics as the most masculine subject and Ormerod showing chemistry. Weinreich-Haste only sampled two single sex schools, whereas Ormerod sampled nineteen schools (single sex and co-educational), which would suggest that Ormerod's placement should be more representative of the total population of schoolchildren. Ormerod refers to the proportion of boys entering the subjects at O level for validation of his 'gender spectrum', but physics attracts more entries from boys than chemistry. The fact that more girls are entered for chemistry than physics suggests that physics is probably the more masculine subject.

"Science is for men". This quote is taken from Evelyn Keller (1978) and was proclaimed by her five-year-old son. Such stereotyping of science as a male activity and consequently of scientists as men is quite common amongst children. In 1957, Mead and Métraux asked 35,000 American secondary school children to complete the sentence "When I think about a scientist, I think of .....". The composite picture which emerged was definitely of a male scientist. This work has recently been repeated in England by Weinreich-Haste (1981). Over a third of the sample specifically referred to the scientist as a man. Tape recordings of 12/13-year-olds talking about science and scientists have also produced a stereotyped view of a male scientist. "They are usually men... well, there's more scope for them and anyway ladies aren't wanted" (Selmes, 1969). "Scientists are boys, and men!" (Lindsay, 1973). In a more structured investigation into stereotypes, Hudson (1968) asked schoolboys to rate a range of science and arts specialists on a number of semantic differential rating scales, including the scale 'feminine-manly'. The scientists were rated as being more manly than the arts specialists. Many studies, mostly American, have shown that the concept of scientist is associated with stereotypically masculine traits (Beardslee & O'Dowd, 1961; Mitroff et al., 1977). Scientists are commonly perceived as being logical, objective, independent, ambitious, competitive, aggressive and unaware of the feelings of others.

2.1.2 Reasons for science's masculine image

The most obvious and widely accepted reason why science has a masculine image is that science is dominated by men. Many figures can be quoted in demonstration of this fact. For example,

(a) The Civil Service Statistics (1979) show that there are 14306 men in science categories, but only 2133 women and they are concentrated in the lowest grade.

(b) Only 0.5% of professional engineers in Britain are women (Chivers, 1981).

(c) Less than 2% of managers, practising research scientists and technologists are women (Curran, 1980).

(d) Male apprentices and trainees at craft level in the electrical and electronic trades outnumber female apprentices 128:1 (Kelly, 1981).

(e) Men form 73% of the labour force manufacturing chemicals, coal

Table 2.1 The number of males for every female studying science at various levels in 1982

Subject	CSE entries	O level entries	A level entries	Undergraduate courses	Postgraduate courses
Physics	4.0	2.7	3.9	6.1	8.2
Chemistry	1.4	1.5	1.9	2.7	4.8
Biology	0.4	0.6	0.7	0.9	1.8

Source: DES, Statistics of Education, School Leavers, CSE and GCE 1982. Universities Statistical Record 1983, University Statistics 1982-1983, Vol. 1 Students and Staff.

and petroleum products (Central Statistical Office, 1983).

Both a consequence and a cause of male domination of scientific occupations is seen in our educational establishments. Males predominate in the physical science subjects throughout the whole of the educational system. As soon as pupils are allowed to choose subjects (at about 14 years), the boys tend to opt for the physical sciences and the girls tend to opt out of them. An HMI report published in 1979 records that roughly 50% of boys study physics and 30% study chemistry in the fourth and fifth years. The corresponding figures for girls are 12% and 18%. In contrast, biology is dominated by girls - 59% of girls study it as opposed to only 32% of boys. Once established, the preponderance of boys in physical science subjects is maintained and indeed amplified, at each successive stage in the educational system. This phenomenon has been extensively documented by writers such as Gardner (1974), Harding (1979), Kelly (1981) and Thompson (1979). Therefore it is not necessary to repeat a detailed analysis here. Table 2.1 indicates the major current trends.

The majority of staff teaching science in educational establishments are also men. Measor (1981) has recorded a girl who said "I think science is a boys subject, most of the science teachers are male, aren't they" and she is correct. 66% of all secondary science teachers are men. 68% of general science teachers are men and in physics and chemistry the proportion rises to 88% and 81% respectively. Biology is the only science subject with more female than male teachers, but with 54% women the numbers are very close (HMI, 1979). As the status and level of teaching posts rises, so the proportion of women occupying them decreases. Only 24% of ILEA schools have a woman as head of their science department (Morrell, 1981). In universities no less than 91% of all full-time teaching and research staff in the science subjects are men (EOC, 1981).

Some writers question whether the domination of science by men is an adequate explanation for science's masculine image. Keller (1978) points out that historically most intellectual and creative endeavours have been executed by men. Yet few of these endeavours, for example, art and writing, are so clearly stamped as male activities. Weinreich-Haste (1979) reminds us that a preponderance of one sex in an occupation does not necessarily result in the occupation being associated with that sex. The two examples that she gives to support her argument are tailoring and electronics assembly work.

Various additional factors have been proposed to explain the obviously close link between science and masculinity. Saraga and Griffiths (1981) suggest that successful scientists display personality traits which are stereotypically masculine and that this enhances the masculine image of science. Other writers maintain that the very nature of science itself, as well as the thought processes and behaviour associated with it, are all intrinsically masculine (Wallsgrove, 1980a; Whyld, 1980). Both scientific and masculine ways of thinking are "analytic, mechanistic, controlling, exploitive, and ultimately destructive" (Fee, 1982). Curran (1980) writes:

> "Scientific knowledge is associated with values which in our society are attributed to men. It is supposed to be objective, rational and impersonal, to the exclusion of feelings, intuitions and emotions, values commonly attributed to women". (p.39)

Wallsgrove (1980b) concurs that science is objective, logical, independent, brave. She also points out that masculine characteristics are valued more than feminine characteristics in our society, and that they are associated with power and control.

Certainly in W. Europe and U.S.A., science is concerned primarily with power, prestige, control, aggression and competition (Harding, 1983, Hinton, 1976). It has been pointed out by Easlea (1981) that practising science enables one group to become dominant, and suggests that men appropriated science to affirm their otherwise fragile masculinity. Saraga and Griffiths (1981) contend that:

> ".....science develops in the service of the dominant interests in any society, both strengthening and defending them. The status of the different sciences, therefore, will vary according to their perceived economic and military significance at any particular time. And the higher their status, then the greater the exclusion of women in male-dominated societies, and the less relevant the subject matter is to issues sanctioned as legitimate for female concern." (p.93)

Hence they argue that the variation in male domination between physics and biology is due to their different historical, economic and military significance, rather than being due to differing gender connotations. Still, in the hierarchy of sciences, the 'hard' sciences at the top of the hierarchy are seen as more masculine than the 'soft' sciences at the bottom.

Science is undoubtedly viewed as masculine, but at the present time the exact causes are still a matter of speculation. Several theories have been proposed. Their usefulness could be assessed by researching the historical development and evolution of the different science subjects' images, and also by analysing current opinions regarding the basis of the masculine image of science.

2.1.3 Reinforcement of science's masculine image

The preponderance of men in science is constantly being presented before the public by the entertainment industry. Films, books, plays, jokes and advertisements invariably portray scientists as men. This practice must help to maintain and strengthen the masculine image of science.

Children are also reminded that science is masculine by their books which usually show people in very traditional, sex stereotyped roles (Weitzman et al., 1972). Elementary reading books continue the conditioning process. Boys are shown playing with scientific toys, men are shown in science-based jobs, but girls and women are shown in caring and nurturant activities (Penrose, 1982; Lobban, 1974).

Science textbooks further reinforce the idea that science is a subject for males. Research into sex bias in textbooks originated in the U.S.A. The illustrations in both primary and secondary level science textbooks were scrutinized for the ratio of male to female figures that appeared (Gaetano, 1966; Heikkinen, 1978). It was found that male figures dominated all the textbooks reviewed. Furthermore, the boys were generally shown in active roles, whilst the girls appeared in passive roles. Recent research in this country has produced similar results. The Leeds Literature Collective (1973) found that most of the pictures and most of the pronouns in primary science textbooks referred to males. At the secondary level, sex bias has been shown in physics (Taylor, 1979; Walford, 1980) and chemistry (Walford, 1981a) textbooks. Both the illustrations and text of these books contain features that favour one sex, i.e. males. The features most commonly mentioned include:

a) male figures appear in illustrations much more frequently than female figures,
b) males are portrayed in active, often science relevant roles, whilst females are shown in more passive, science irrelevant roles (Walford, 1980),
c) occupations are usually sex role stereotyped, e.g. engineers are men, nurses are women (Taylor, 1979),
d) when people are mentioned in the text they are generally male,
e) the pronoun 'he' is used much more frequently than 'she', particularly when referring to a scientist or technologist (Samuel, 1981),
f) examples and analogies are based upon objects and activities ~~are based upon objects and activities~~ stereotypically associated with males (Taylor, 1979),
g) questions and problems generally relate to male figures and male interests (Kelly, 1976b; Walford, 1981b).

Recently published physics textbooks show slightly less sex bias than those published in the late 60s and early 70s, but this is largely due to fewer people being shown in illustrations and mentioned in questions (Walford, 1981b). Where people do appear in illustrations in O level and CSE texts, there are, on average, about five times more men than women, so the position of females is still very marginal. Even the 'Science in Society' project materials produced by ASE (1981) perpetuate the link between science and 'men's society' (Fawcett Society, 1981). Instead of including ample examples of the applications of science and its social relevance for both sexes, the pupils' readers concentrate upon only one sex. Harding (1981) charges that "they were written by men, from a peculiarly male perspective on life, with boys in mind". The Fawcett Society (1981) suggest that the involvement of more women at the development stage might have resulted in a project which acknowledged the presence of women in society. This point applies to all material produced for science education. Until a higher proportion of women are involved in the planning and writing of educational materials, the presentation of science as a subject for boys, but not girls, is likely to continue.

The 'male oriented' bias of science is also maintained in other subtle ways. Posters depicting great scientists and their scientific discoveries constantly remind pupils that historically science has been a male activity (Kelly, 1982). Even the language used to describe their work often invites comparison between intellectual creativity and male sexual activity (BSSRS, 1975). For example:

> "Fellows of the Institute of Physics probe into the dark, have penetrating insights, and make breakthroughs".
>
> (Preece, 1977)

A recent study of the textbooks used on method courses by American students training to teach science at the primary level revealed the following:

(a) roughly two-thirds of the illustrations showed male figures,

(b) over seven times as much content space was allocated to males as to females,

(c) practically no mention was made of sex differences (Sadker & Sadker, 1979).

The rapidly growing body of literature concerned with the detrimental effects of sex differences, sex stereotyping and sexism in science education does not yet appear to have been incorporated into teacher training programmes. Such is also the case in this country (Kelly, 1981; Whyte, 1981). Thus new teachers are entering the teaching profession ill equipped to recognise, let alone try to counteract, the many factors that help to convey the impression that science is a masculine subject.

Teachers, through the best of intentions, frequently reinforce the idea that science is a subject for boys. In attempting to make science relevant to their pupils, most teachers quote examples from everyday life. Unfortunately, like the textbooks, they often pick examples (e.g. guns, engines, vehicles) that are of much greater interest to boys than to girls. In fact, girls may have little previous knowledge or experience of such devices. But, as Spender (1982) writes, "The male experience becomes the classroom experience" (p.59).

2.1.4 Summary

1. School children and students believe that physical science has a masculine image. When asked to rate the physical science subjects on a masculine-feminine semantic differential rating scale, their ratings are towards the masculine pole. A high proportion also agree that physical science subjects are more suitable for boys than for girls, and that they are better done by boys than by girls. Since science is stereotyped as a male activity, school children stereotypically

think of scientists as being men. No studies were located that investigated the opinions of adults in this country regarding the masculine image of science. Thus we do not know how science teachers view the gender connotations of the subjects that they teach. Neither do we know the sex that they associate with the occupation of scientist.

2. A number of factors are believed to contribute to the masculine image of science.

(a) Science is dominated by men. Most science students, most science teachers, most researchers in science, and most people working in science related occupations are men.

(b) Scientists are associated with stereotypically masculine traits. For example, they are commonly perceived as being ambitious, competitive, aggressive.

(c) The cognitive styles associated with science are stereotypically masculine ways of thinking, and are described by adjectives such as objective, logical, analytic. Some writers maintain that science itself displays such characteristics.

(d) The very content of science is intrinsically masculine. Science is concerned with prestige, power, and control.

3. The close association between science and males is maintained and strengthened by:

(a) the entertainment industry - adverts, films, comics,

(b) children's reading books,

(c) pupils' science textbooks, teacher training textbooks,

(d) the analogies and examples used by science teachers.

2.2 SEX STEREOTYPES

This section is divided into two parts. In the first part a definition of 'sex stereotype' is presented and commented upon. Then some of the traits stereotypically linked with each sex are recorded, and a few examples of the ways in which sex role stereotypes can influence people's perceptions and judgements are discussed. Although only one study involved teachers, it can reasonably be argued that the findings and conclusions are nevertheless generally applicable to teachers. Teachers are not isolated from the rest of society and so their beliefs, perceptions and behaviours are unlikely to be very different from those of the general public. The second part of the section refers specifically to school settings and the sex stereotypes held by teachers.

2.2.1 General aspects of sex stereotyping

Sex stereotyping has been defined in the Rules and Regulations of the 'Women's Educational Equity Act' (United States Office of Education, 1976) as "the attribution of behaviors, abilities, interests, values and roles to a person or group of persons on the basis of their sex". Sex stereotypes reflect over-simplified ideas about men and women. They ignore individual differences and assume that all individuals of the same sex share the same abilities, interests and aspirations. For example, the belief that all boys are interested in science and most boys are good at science is a sex stereotype.

Sex stereotypes are acquired and maintained by socialization processes (Open University, 1982). Children usually first encounter sex stereotyping in their homes. Initial stereotypes are later reinforced and extended by a variety of agencies, e.g. the school, the community, the mass media. Gradually sex stereotyped beliefs are incorporated into the self-concepts of both sexes. Their adoption not only constrains people's perceptions of themselves and others, but also influences their actions, since sex stereotypes prescribe appropriate behaviour for each sex.

Sex stereotypes are well documented in the psychological literature (Broverman et al., 1972; Fernberger, 1948; Lunneborg, 1970; Rosenkrantz et al., 1968). According to this research, men are popularly believed to be more dominant, independent, competitive, intellectual, athletic, unemotional, self-confident, ambitious, aggressive, decisive, logical, analytical and objective than women. On the other hand, women are commonly viewed as being more submissive, dependent, emotional, excitable, irrational, conforming, affectionate, kind, sensitive, warm, sympathetic, understanding, gentle and nurturant than men. These stereotypes appear to have remained relatively stable over the three decades spanned by the studies cited. Furthermore, they are shared by people differing in sex, age, marital status and education (Broverman et al., 1972). Although most of the research involved American samples, a recent study by Williams et al. (1977) indicated that very similar sex role stereotypes prevail in both the United States and England.

Numerous studies have shown that both men and women attribute greater social value to masculine traits than to feminine traits (Kitay, 1940; McKee & Sherriffs, 1957; Rosenkrantz et al., 1968). The frequent devaluing of stereotypically feminine attributes results in females being largely associated with traits which are commonly perceived as inferior or negative. Such views are not only held by the general population, but also by professionals. Broverman et al. (1970) found that clinicians (clinically trained psychologists, psychiatrists, and social workers) judged the personality and behavioural traits of a mature, healthy male to be closer to ideal standards of mental health than were the traits of a mature, healthy female. These findings have been repeated with male, but not female, counsellors-in-training (Maslin & Davis, 1975). Garman and Plant (1974) conducted a similar investigation with teachers and found that they also perceived the attributes which characterize competency for adult males to be significantly closer to those attributes ideally possessed by a healthy, mature, socially competent adult than are the attributes which characterize competency for adult females. Again, differences appeared between male and female subjects. Female teachers judged female adults to be more like the general standard of maturity and social competence than did male teachers.

Sex role stereotypes have also been shown to influence the evaluation of males and females in managerial positions (Rosen & Jerdee, 1973, 1974a). Successful middle managers are thought to possess characteristics, attitudes and a temperament that are more commonly associated with men than with women (Schein, 1973, 1975). Sex role stereotyping can result in women being perceived as less qualified than men for management positions. Rosen and Jerdee (1974a) demonstrated a tendency for male administrators to discriminate against women when considering supervision, career development and promotion.

Research in other employment areas indicates interaction between the sex typing of an occupation and the sex of an applicant (Cash et al., 1977; Ward, 1977). For example, Cohen and Bunker (1975) found that university recruiters recommended more female than male applicants for a feminine occupation, while the reverse was true for a masculine occupation. Ward (1981b) suggests that the sex appropriateness of an occupation, its prestige and the competence of the judge all play a part in the differential evaluation of males and females.

2.2.2 Sex stereotyping in schools

Incidents of sex stereotyping are frequently encountered in connection with education. It is now well established that children's literature presents and thus helps to perpetuate very sex stereotyped roles and behaviours (Lobban, 1974; Stones, 1983; Weitzman et al., 1972). More recently it has been demonstrated that neither the manner of presentation nor the actual content of many school textbooks it totally sex neutral (Sadker & Sadker, 1982; Scott, 1980). Even in subjects which are very abstract and impersonal, such as maths and science, the textbooks still convey messages about appropriate sex roles to pupils (Taylor, 1979; Walford, 1980). Out of school, in more informal learning situations, sex stereotyping is further reinforced by the mass media (MacDonald, 1981; Manstead & McCulloch, 1981; Sutherland, 1981), and the language styles in common use (Miller & Swift, 1979; Spender, 1980).

A number of workers have studied and documented the many subtle ways in which the organisation of schools, as well as the procedures and processes that take place within schools, convey sex stereotyped ideas to pupils and continually reinforce them for both pupils and staff (Delamont, 1980; Whyld, 1983). The reports indicate that schools foster traditional sex stereotyped conduct and beliefs, and that teachers are instrumental to the process.

Teachers are quite familiar with the sex stereotyped beliefs held by society at large. A group of teachers (from primary, secondary and tertiary establishments) indicated to Zimet and Zimet (1977) that males are perceived by the general public to be significantly more achievement oriented, autonomous and aggressive than females. On the other hand, they indicated that females are viewed as being significantly more deferent than males.

Teachers believe that their own attitudes to sex roles are more liberal than those of the general public (Evans, 1982). However, Evans (1982) reports that teachers' perceptions of male and female roles are relatively conservative. Ricks and Pyke (1973) are even more disparaging. They liken female teachers' attitudes towards women's liberation with those of suburban housewives. The fact that male teachers tend to have less positive attitudes towards issues of working mothers (Ricks & Pyke, 1973; Tetenbaum et al., 1981), simply highlights further the conservative nature of teachers' attitudes towards sex roles.

The traditional views of sex roles held by teachers appears to influence their expectations and perceptions of their pupils. Frazier and Sadker (1973) and Sharpe (1976) report studies in which secondary teachers were asked to select adjectives to describe what an adolescent girl or boy should be like. In both studies the replies reproduced stereotypes of the male and female roles. Adjectives such as submissive, dependent, emotional, conscientious and obliging were used for the girl, whilst the boy was described by adjectives such as aggressive, independent, assertive, inventive and active. As in other studies, the characteristics associated with the boy are invested with greater social desirability.

Once teachers link certain qualities with one sex, then they may perceive pupils of that sex to be more suited for the study of subjects which demand those qualities. Thus boys being objective, inventive and unemotional should be well suited for science, maths and technical subjects. Girls are more subjective, emotional and less curious and so are more likely to be interested in literature or domestic subjects (Whyte, 1981). Such reasoning extends beyond the sex typing of subjects to the sex typing of topics within a subject. Wolpe (1977) described an English master who expected different kinds of written work from boys and girls. He would set a number of essay titles and explicitly state which were suitable for boys and which for girls.

Teachers are very aware of behavioural differences between boys and girls. 73% of the teachers in a study conducted by Ricks and Pyke (1973) believed that boys and girls behave differently. Furthermore, teachers can readily identify behaviours that distinguish between the sexes. Boys are seen as active, restless, noisy and boisterous, whilst girls are perceived to be passive, conformist, obedient and orderly (Clarricoates, 1980); Davies & Meighan, 1975; Ricks & Pyke, 1973). The degree to which teachers' perceptions coincide with actual pupil behaviour is not altogether clear, since most studies of teachers' beliefs and pupils' classroom behaviour has been conducted separately. However, Clarricoates (1980) points out that the readiness with which teachers can produce lists of stereotyped behaviour indicates that teachers frequently classify pupils according to their sex. Of even greater concern is Davies and Meighan's (1975) finding that teachers place different interpretations upon stereotyped patterns of behaviour by each sex. Misbehaviour by boys is regarded as 'prank-playing', but when girls misbehave they are 'devious' and 'insolent'.

As a result of stereotyped beliefs about sex differences between boys and girls, teachers behave differently towards boys and girls and hold different expectations of them (Brophy & Good, 1974; Davies & Meighan, 1975; Sears & Feldman, 1974; Sharpe, 1976). In many cases this differential behaviour is neither unintentional nor unrecognised. Sears and Feldman (1974) reported that about half of their sample agreed that they did behave differently to boys and girls. Just over half of Ricks and Pyke's (1973) sample were of the opinion that male and female pupils expect different treatment. Although both male and female teachers think that they do, and should behave differently towards boys and girls, Ricks and Pyke suggest that the behaviour of male teachers may be more effective in conveying traditional sex role prescriptions, since male teachers are more likely to believe that girls want chivalrous treatment.

For many years teachers have been inclined to believe that boys and girls possess different work related attributes. In 1923, the Board of Education reported that many teachers thought that boys were generally more original, constructive, experimental and logical than girls. Moreover, boys were judged to be more analytical and to be better able to understand and apply general principles. In contrast, girls were seen to be more intuitive, persevering, industrious, imitative, patient, conscientious and neat. Consequently, the teachers thought that boys sought self-expression in 'investigation and construction', whilst girls used 'artistic and emotional channels'. According to a more recent report, teachers still think of boys as being more logical and quicker to grasp new concepts, whilst girls are seen as being conscientious, precise and better at written work (Davies & Meighan, 1975). In the same study, 72% of the teachers (both male and female), in response to a forced choice question, said they would prefer to teach boys. Their reasons included the attributes listed above. Ricks and Pyke (1973) also found that teachers who preferrred to teach boys gave similar reasons for their preference. The only reason mentioned for preferring to teach girls was that girls are easier to discipline. Such findings clearly show that the qualities and attributes that are associated with male pupils are the ones that teachers value most. As a result, girls tend to receive less attention from teachers and to occupy a very peripheral position in classroom life (Stanworth, 1981).

Sex differentiated perceptions of pupils' personality traits, scholastic aptitudes and interests lead to teachers holding different expectations for boys and for girls. During the years of schooling, especially at the secondary level, teachers frequently expect higher academic standards from boys (Frazier & Sadker, 1973). After school, teachers expect the adult lives of boys and girls to be very different (Delamont, 1980). Stanworth (1981) found that marriage and parenthood featured prominently in teachers' visions of the futures of their female pupils, but hardly at all in boys' futures. Boys were seen in jobs involving responsibility and authority. When girls were envisaged in paid employment it was always of a traditional nature, e.g. secretarial work, nursing, teaching. However, two-thirds of the male teachers interviewed could not visualize their female pupils in any occupation at all when they left school. Such findings clearly show that teachers expect the female adult role to revolve around housework and childcare. If a woman works, the job should be of a subordinate and nurturant nature.

Most of the studies referred to above involved non-science teachers. Very few studies have concentrated specifically upon the perceptions and attitudes of science teachers. However, Seale et al. (1982) suggest that science and craft teachers have quite clearly sex stereotyped views, and Delamont (1982) has provided some sex stereotyped teaching exchanges that were encountered in science lessons. Thus it appears that science teachers' sex role perceptions are no less traditional than those of teachers of other subjects, and may, in fact, be even more traditional.

Teachers who hold sex stereotyped beliefs about what are appropriate behaviours, abilities, personality traits and careers for girls and for boys are likely to convey their beliefs to their pupils and thus to influence the pupils' behaviour, achievements and development (see section 2.4). This could result in pupils limiting their options and restricting their aspirations and potential. Rosenthal and Jacobson (1968) have shown that teachers' expectations can affect their pupils' classroom behaviour and academic achievement. One way in which expectations are undoubtedly conveyed is via teacher-student interactions (mentioned above).

Sex role stereotypes are not static. There have been changes over the last couple of decades, probably as a result of the dissemination of ideas from the women's movement. Evans (1982) found that a sample of Australian teachers had perceived changes to adult sex roles, especially for women. It is interesting that the male teachers saw the changes as more marked and significant than did the female teachers. Changes in adult sex roles appear to have influenced the sex roles and behaviours displayed by children. The majority of teachers believe there have been changes, especially with girls. They perceive that girls are becoming more aggressive, assertive, argumentative and active (Evans, 1982; Ricks & Pyke, 1973). However, only a minority of teachers feel that they should facilitate this change. Ricks and Pyke (1973) found that 57% of a sample of American teachers thought that it is not a teacher's responsibility to influence children's attitudes towards sex roles. It may thus be concluded that there is no concerted effort by teachers to transmit anything other than the traditional sex role stereotypes. Instead of modifying sex stereotypes in schools, the teachers' traditional beliefs help to maintain and perpetuate sex roles.

2.2.3 Summary

1. Sex role stereotypes reflect over-simplified ideas about men and women, by assuming that all individuals of the same sex share the same abilities, interests and aspirations.

2. Men are commonly viewed as being more self-confident, ambitious, aggressive, unemotional, intellectual, objective; whereas women are more submissive, irrational, sensitive, gentle and nurturant.

3. Teachers hold relatively conservative sex stereotyped beliefs. These beliefs help to maintain and perpetuate traditional sex stereotyped views, expectations and conduct in schools.

4. The traditional views of sex roles held by teachers influence their perceptions of their pupils.

(a) Teachers stereotype the behaviour of their pupils. Boys are perceived to be more active, noisy; whereas girls are more obedient, orderly.

(b) Teachers stereotype the personality characteristics of their pupils. Boys are perceived to be more independent, assertive, inventive; whereas girls are more subjective, emotional, dependent.

(c) Teachers stereotype pupils' work related attributes. Boys are perceived to be more logical, experimental, analytical; whereas girls are more conscientious, industrious, neat.

Teachers value the qualities and attributes associated with male pupils more than those associated with female pupils.

5. Sex differentiated perceptions of pupils' personality traits and scholastic aptitudes lead to teachers holding different expectations for boys and for girls. Teachers frequently expect higher academic standards from boys.

6. Teachers stereotyped beliefs about sex differences causes them to behave differently towards girls and boys.

7. Very few studies have concentrated specifically upon the sex stereotyped beliefs and sex differentiated perceptions of science teachers. However, there are indications that science teachers hold relatively traditional sex stereotyped beliefs.

ATTRIBUTION PATTERNS

By adulthood, all people have acquired a set of beliefs or assumptions that determine their understanding of and influence their behaviour in different situations. One component of this belief system that has received considerable attention from psychologists concerns the causal relationships used to explain various outcomes. Research has shown that a person's attributions or beliefs about the causes of success and failure not only help to determine their performance in achievement settings, but also influence their future expectations of success and subsequent achievement strivings. In addition, attributions have a strong effect on how a person reacts to their own and other people's successes and failures (Frieze, 1980; Weiner, 1974, 1979).

People use a variety of causes or attributions in explaining a particular success or failure (Frieze, 1976; Weiner, 1974). However, most studies of causal judgements in achievement-related contexts have employed structured methods and have concentrated upon just four factors: ability, effort, task difficulty and luck. Success is generally explained in terms of a person having high ability, trying hard, having good luck, and/or that the task was relatively easy. Failure may be seen to result from low ability, lack of effort, bad luck, and/or the task being difficult. Research using an unstructured, open-ended approach has indicated that additional causal factors are frequently used to explain personal successes and failures, as well as those of others. These include fatigue, illness, mood, teacher bias (Frieze, 1976).

Causal attributions can be classified along three dimensions: locus, stability and controllability (Weiner, 1979). Locus, proposed by Rotter (1966), refers to the location of a cause, i.e. internal or external to a person. Ability and effort are considered to be internal factors because they originate within the person, whereas task difficulty and luck are external factors since they arise from environmental sources. The locus dimension has been shown to influence the affective reactions of pride and shame (Weiner et al,, 1978). The second dimension differentiates causal elements in terms of their stability over time. Ability and task difficulty are perceived as relatively stable causes, but effort and luck may be highly changeable. The stability dimension relates to expectancies of future success and failure (McMahan, 1973; Valle & Frieze, 1976). The third dimension, controllability, refers to a person's perceived voluntary control over a cause. Causes such as ability and luck are not perceived as being under personal control, but effort is controllable. The controllability dimension relates to sentiments and evaluations of others (Weiner, 1979).

The stability of causal attributions, together with the level of initial expectation of success and the level of performance achieved on a task, act together to maintain expectancies for future outcomes at a similar level. An expected outcome at a task, either good or poor, is usually attributed to stable factors such as ability (Feather, 1969; Feather & Simon, 1971). Consequently, expectancies for future performances remain unchanged. If an outcome is unexpectedly high or low, it will tend to be attributed to unstable factors such as luck. Unstable causes, by definition, suggest that the present outcome is atypical and may not be repeated again. Hence expectations for future performances remain the same. This leads to a self-fulfilling prophecy where those who expect to succeed maintain their high expectations, and those with low expectations do not change them regardless of their actual level of performance. Freize (1980) suggests that this effect not only operates for an individual, but also when a person makes attributions about another person. Thus the causal attributions used by a teacher to explain the success or failure of a pupil probably reinforce that teacher's expectations of the pupil.

2.3.1 Sex differences in causal attributions

There is some evidence that males and females tend to make different attributions for their performance on achievement tasks. Males more often than females have been found to use ability to explain their successes. Females have been found to underestimate their level of ability and overestimate the contribution of luck to their performance, i.e. to employ more external attributions (Bar-Tal &

Frieze, 1977; Deaux & Farris, 1977).

Recent work has questionned the existence, pervasiveness, magnitude or interpretation of sex differences in self-attributions (McHugh et al., 1982). Travis (1982) looked for sex differences at different points in the attributional model and failed to find any sex differences in subjective evaluations of success, causal attributions, or expectations for future performance. Other workers have produced predicted results, but have then interpreted them from a different perspective and so have arrived at alternative conclusions (Sweeney et al., 1982).

The technique of meta-analysis has also been used to reassess the influence of a person's sex upon their achievement self-attributions. An analysis carried out by Frieze et al. (1982) showed that although men made stronger ability attributions than women when causal attributions were inferred from stated levels of ability possessed and men attributed their successes and failures less to luck than did women, there were no strongly supported sex differences in attributions. Another meta-analysis by Sohn (1982) demonstrated that even in studies where sex differences did occur, these differences accounted for less than 5% of the variance. The only exception concerned luck attributions. Women tended to use luck to explain successful outcomes more than men did.

These studies make clear the contradictory nature and lack of consequential relationships between sex and attribution behaviour. It could be that sex differences in causal attributions have been created by publication biases, i.e. the tendency to publish only those studies which reject the null hypothesis (Greenwald, 1975), and that they do not in reality exist. Alternatively, the inconsistencies in the attribution literature could result from failure to adequately consider various situational and dispositional determinants of sex differences in attribution (McHugh et al., 1982). If people are asked to explain the performance of a stimulus person rather than their own performance, then sex differences in the attributions are usually detected. When men are presented in stimulus material as succeeding, their success is attributed more to skill, whereas equivalent performances by women are attributed more to luck (Deaux & Emswiller, 1974). On the other hand, causal attributions for failure display a different pattern. Men's failure is seen as having been caused by bad luck, women's failure by lack of ability (Feather & Simon, 1975). Thus the success of women is attributed to unstable external causes, whilst their failure is attributed to stable internal causes. The pattern for men is the exact opposite.

In these studies, it is implied that sex role attitudes and expectations mediate in the attribution process. Individuals interpret and explain the performance of a stimulus person using their own sets of beliefs about the sexes. Undoubtedly sex stereotyped ideas about the innate abilities and aptitudes of men and women must influence judgements. Beliefs about the appropriateness of the task for each sex, the likelihood of each sex possessing the necessary skills, and the possible consequences of success and failure for each sex may also bias evaluations and judgements.

The nature of the achievement task itself should not be overlooked, for it has been shown to exert an important impact on causal attributions (Frieze & Snyder, 1980). For example, sex differences in attribution patterns are most marked for sex typed achievement tasks. This occurs whether people account for their own performance (Gitelson et al., 1982) or the performance of others (Feldman-Summers & Kiesler, 1974). Success on a sex-inappropriate task or failure on a sex-appropriate task is an unexpected outcome and therefore is likely to be attributed to unstable or external causes, whereas failure on a sex-inappropriate task or success on a sex-appropriate task is expected and so is likely to be attributed to stable internal factors (Etaugh & Brown, 1975). Thus Deaux (1976) found that successful performance by men on a masculine task involving mechanical objects was attributed to skill, while equivalent performance on the same task by a woman was more likely to be attributed to luck. When the task involved feminine objects (kitchen utensils), the men were credited with less ability and the women with more, but ability was still perceived to be a more important explanatory factor for male performance than for female performance. Similar findings are reported by Deaux and Farris (1977). Expectancies, evaluations and attributions to ability tended to be more equal on a feminine task, but not reversed in favour of females. These findings are consistent with beliefs and expectations that men generally have more ability than women (Deaux, 1976). This explanation is further supported by the work of Bond and Deming (1982). They identified biases relating to the sex of the stimulus person in explanations for failure, which suggested that failure was an anticipated outcome for women.

2.3.2 Attributions in science lessons

If the attribution patterns described above are also displayed by schoolchildren, then they would be expected to give sex differentiated attribution patterns when accounting for their successes and failures at school, especially in sex typed subjects. Physics and chemistry are commonly accepted as being very masculine subjects, and their study as being more appropriate for boys than for girls (see section 2.1.1). Thus the success of girls in these subjects is likely to be viewed as being due to unstable external factors, and their failure to stable internal factors, e.g. lack of ability. Such attributions would also help to resolve conflict that girls may experience if they accept the sex stereotyped view that girls have a natural inferiority in science subjects. If girls explain their success in science as being due to luck or ease of task, then they can still maintain their belief that they have little aptitude for science.

Physics and chemistry are considered to be difficult subjects. It is generally found that more difficult tasks tend to produce more ability attributions for success and failure (Frieze & Snyder, 1980). Therefore a relatively high proportion of ability attributions would be expected for the physical science subjects. Since females are more likely to make ability attributions for failure than males, the liklihood of girls explaining their failure at science in terms of their lack of ability must be further increased. The fact that many girls may have poor initial estimations of their aptitude for science would not help the situation.

Pupils' attributions of their successes and failures in science have been little studied. However, similar theories regarding sex differentiated attribution patterns in maths have been proposed and tested (Frieze, 1980). Dornbusch (quoted by Beckwith & Durkin, 1981) found that more girls than boys accounted for a poor maths grade in terms of lack of ability. Lorenz (1982) reports that when pupils were asked to explain their maths grade, the girls believed that they had lower ability for maths than did the boys. On a spatial task Gitelson et al. (1982) found that girls attributed to themselves less ability and saw the task as being more difficult than did boys. However, the work of Parsons et al. (1982) indicates that the type of questions asked (open-ended or rank-ordered) can influence the results obtained. When two different attribution measures were used, consistent sex differences in attributing failure to lack of ability did not emerge.

Children who attribute failure to lack of ability or other factors that they cannot control, may suffer from 'learned helplessness'. This condition exists when failure is perceived as inevitable and insurmountable (Dweck & Reppucci, 1973). It is educationally dysfunctional since children who exhibit learned helplessness believe that they will continue to fail in the future and this adversely affects their effort and subsequent performance. In contrast, children who attribute failure to lack of effort, or other controllable factors, tend to see failure as a temporary setback which can be changed to success through greater effort. Thus effort attributions are likely to have beneficial effects upon future performance. In accord with the findings described above, girls are more likely than boys to exhibit learned helplessness (Dweck et al., 1978; Dweck, 1980).

2.3.3 <u>Summary</u>

1. The success of men in frequently attributed to skill, whereas the success of women is often attributed to luck. In contrast, men's failure is seen as having been caused by bad luck, women's failure by lack of ability. Sex differences in attribution patterns are most marked for sex typed tasks.

2. In maths, more girls than boys attribute a poor grade to lack of ability.

3. The attributions that pupils make for their successes and failures in science have not been studied. However, since physics and chemistry are viewed as masculine subjects and difficult subjects, it is to be expected that girls would attribute their success to luck and failure to lack of ability.

4. No studies were located of the attributions that science teachers make to explain the success and failure of their male and female pupils. These attributions are important since they probably

reinforce the expectations that a teacher holds of a pupil.

2.4 TEACHER EXPECTATIONS

Expectations for the behaviour of others is a basic component of all social relationships. Within education, the focus of attention has been directed towards the expectations of teachers concerning their pupils. The work on teachers' expectations has encompassed two broad areas, those of pupils' behaviour and their academic achievement. Although often overlapping, as indicated by halo effects, these two areas have often been separated for detailed investigation. The following review looks at recent research relating to teacher expectation and pupil performance.

Teacher expectation has been the subject of much research and many review articles (e.g. Braun, 1976; Burstall, 1978; Dusek, 1975) because it is seen as a variable of educability of enormous potential importance. In fact, proponents of teacher expectation claim that it is the major variable of educability (Nash, 1973). They hypothesise that "the teacher for varied reasons perceives competencies and potentialities of children differently and that these expectancies are reflected in his interactions with children to produce differential performance among learners, thus fulfilling his prophecy" (Braun, 1976).

'Self-fulfilling prophecy', 'teacher expectation' and 'teacher faith' are the terms that have been coined to signify this tendency for the teacher to create a reality corresponding to his perceptions. As Thomas wrote in 1928 "If men define situations as real, they are real in their consequences". Some years later Merton (1968) suggested that the self-fulfilling prophecy is "in the beginning, a false definition of the situation evoking a new behavior which makes the originally false conception come true". However, this definition is restricted to situations involving an original misperception. The definition offered by Rosenthal and Jacobson (1968) "How one person's expectation for another person's behaviour can quite unwittingly become a more accurate prediction simply for its having been made" has much wider application.

2.4.1 Empirical studies

The impact of interpersonal expectations was first demonstrated in psychological research. Rosenthal (1976) has reported a series of psychological experiments that he conducted which all indicate that one person's expectation of another's behaviour may serve as a self-fulfilling prophecy. Within a laboratory situation, Rosenthal has demonstrated that the expectancy of the behavioural scientist can significantly influence the outcome of experiments. This happens even when the subject cannot see the researcher. And, even more surprisingly, it occurs when the subject is not human but a rat (Rosenthal & Lawson, 1964).

Research into expectancy effects swiftly spread to educational settings. In 1963 Clark had argued that the generally poor performance of American ghetto children might be due to low teacher expectations which became self-fulfilling prophecies. A controversial study by Rosenthal and Jacobson (1968) provided empirical support for the idea. They had administered a little known non-language intelligence test, misrepresented as 'The Harvard Test of Inflected Acquisition', to primary children in grades 1 - 6 (6-12 years). The teachers had been told that the test identified children who were likely to show marked intellectual improvement within the year. The investigators then arbitarily designated a randomly selected 20% of the pupils as being intellectually 'about to bloom'. Subsequent testings took place four, eight and twenty months later. The findings of the study are complicated because they depend upon what combinations of grade, testing time, sex, ability grouping and intelligence sub-test are considered. However, Rosenthal and Jacobson concluded that their data indicated that if teachers expected intellectual blooming in specific children, such gains would, in fact, result.

'Pygmalion in the Classroom', the book in which Rosenthal and Jacobson reported their study, aroused immediate interest in both academic and lay circles. The work was heavily criticized on methodological grounds and the conclusions were challenged by many researchers. Major criticisms include poorly defined sampling procedures, inadequate data analysis and misleading graphs and tables (Jensen, 1969). Claiborn (1969) points out that conclusions were based upon simple gain scores which were not corrected for known pre-test differences, and which may well have been partly attributable to regression effects. A procedural criticism is that the test was administered by the class teachers (Jensen, 1969; Snow, 1969). Furthermore, the test which was used, the Test of General Ability, is considered inadequate for young children of low socio-economic status (Thorndike, 1968). It gave an average IQ of 58 for the first grade pupils. Such a low score is highly suspect. But probably the most striking point is that the teachers did not remember the names of the 'bloomers' (Braun, 1976). Braun concluded that "biased reporting, magnification of selected findings, and over-dramatization of conclusions raise serious questions ...". Thorndike (1968) was even more scathing in his assessment of the study: "It is so defective technically that one can only regret that it ever got beyond the eyes of the original investigators!"

The value of Rosenthal and Jacobson's work is that it was a pioneer study, which stimulated much further research into the possible effects of teacher expectation on pupil achievement. Claiborn (1969) attempted to replicate Rosenthal and Jacobson's study, but over a shorter period of time - two months. Although a substantial overall increase in IQ was noted, there was no significant differential change between 'bloomers' and controls. Fleming and Anttonen's (1971) study also failed to show expectancy effects. This was a large scale study involving thirty nine teachers and over one thousand seven year old children. The children were tested and then their teachers were given either PMA (Primary Mental Abilities) percentages, traditional IQ scores, IQ scores inflated by 16 points, or no IQ scores. Upon testing at the end of the study, the children with inflated IQ scores did not show greater relative gains. Numerous other studies have also failed to support the findings of Rosenthal and Jacobson's original study (e.g. Dusek & O'Connell, 1973; Fielder et al., 1971; Jose & Cody, 1971; O'Connell et al., 1974).

However, in spite of this disappointing and discredited beginning, research on the expectation phenomenon has continued; probably because it appears to be psychologically and philosophically logical. Also investigations have been sustained by the encouragement of a number of studies that lend convincing, if not unequivocal, support to the expectancy hypothesis. Rosenthal (1973) noted that by 1973, 84 out of 242 studies supported the notion that teachers' (or experimenters') expectations do affect pupils' (or laboratory volunteers') performances.

An investigation by Pippert (quoted by Braun, 1976) casts doubt upon the widely used technique of artificially creating expectancies. Pippert found that pupils of teachers who doubted the stated purpose of the experiment gained substantially less than pupils of teachers who had no doubts. Mendels and Flanders (1973) suggest that contrived sources may be much less powerful determinants of expectations than natural factors. Certainly it is noticeable that studies investigating teachers' attitudes and expectations which have been built up naturally, rather than as a result of contrived information, tend to yield more positive results.

Palardy (1969) compared the reading performance of first grade pupils (6 years) taught by teachers who believed that girls learn reading more easily than boys with those taught by teachers who did not share that belief. The teachers' preconceptions did seem to constitute a self-fulfilling prophecy: girls were superior in the classrooms of teachers who expected it, but not in those of teachers who did not. A large scale study on 'educationally deprived children' (US Office of Education 1970, quoted by Rosenshine, 1971) also reported "an extraordinarily consistent relationship between teacher expectations and the reading achievement gains of pupils".

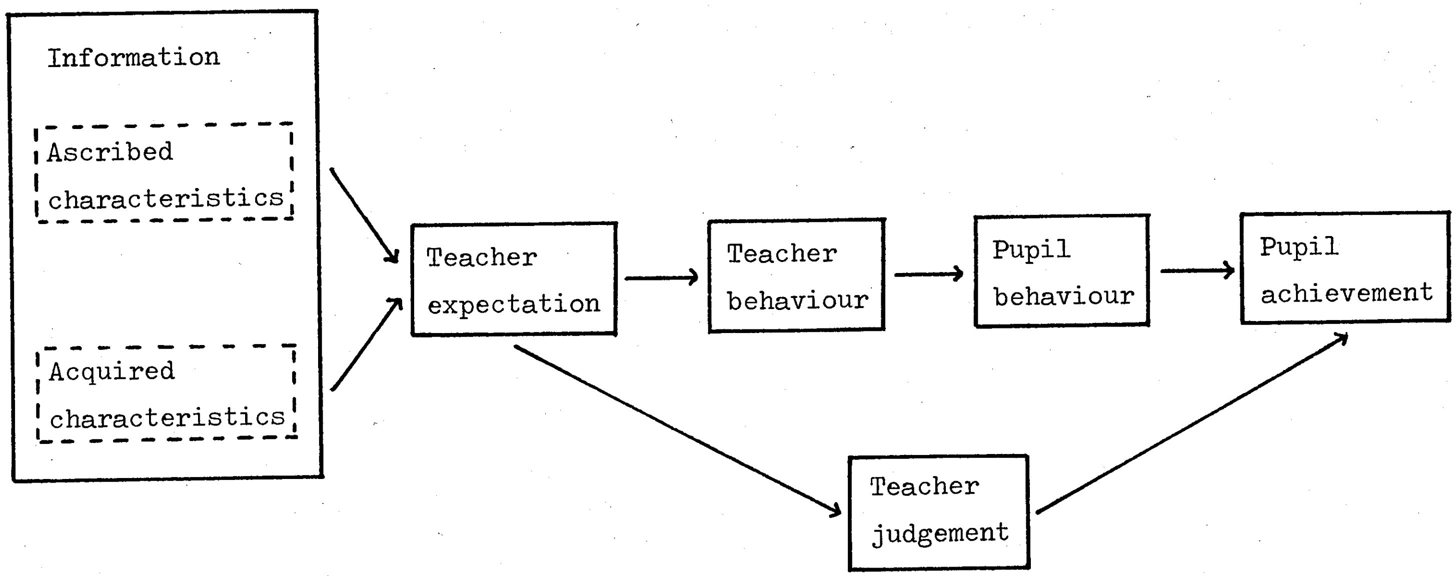
Burstall (1968) working with the NFER studied the attitudes of teachers towards the teaching of French to low ability pupils. It was found that pupils of low general ability, who nevertheless scored above the mean for their particular group, were concentrated in schools where teachers expressed favourable attitudes. In these schools the head also had a favourable attitude towards the teaching of French to children of all levels of ability.

The above studies indicate the existence of a close association between the teacher's attitudes and expectations and the pupils' achievement. Exactly how such self-fulfilling prophecies operate is uncertain, as most studies have tended to consider teacher expectation at a descriptive rather than explanatory level.

2.4.2 The mechanism of teacher expectancy effects

Models could provide useful research frameworks for conceptualizing the phenomenon and guiding future research into the mechanisms which underlie teacher expectancy effects. Such models have been

Figure 2.1 A model of the teacher expectation hypothesis



proposed by Braun, 1976; Brophy and Good, 1970; Persell, 1977; West and Anderson, 1976. An amalgamation is presented in figure 2.1 This model is used to structure the following brief review of the factors and mechanisms believed to be important in the formation and conveyance of teacher expectations.

It is proposed that information relating to a pupil generates certain expectations on the part of the teacher for that pupil. The information can either concern attributes which a child possesses on entering school (ascribed characteristics such as sex, race, perceived social class) or attributes which a child acquires during the educational process (acquired characteristics such as academic stream or set, school reports). Both sources of information may be subjective or objective in nature. Teacher expectations may lead, under some circumstances, to biased judgements of pupil behaviour and performance, and/or differential teacher behaviour.

Systematic investigation of the teacher expectation-teacher behaviour link has been undertaken by a number of researchers and their findings provide evidence of both quantitative and qualitative correlations between teachers' expectations and behaviour. Teachers apparently devote more time and attention to high achievers than low achievers, and their pupils notice what is happening. "I think the teachers could have spent time with the ones who didn't know so much but no, they just looked after those people who already knew things" complained one school leaver (Collaborative Research, 1977). It seems that it is the well endowed children who get the most teacher contact, while teachers believe the reverse is (and should be) taking place. Burstall (1978) has labelled this occurrence 'The Matthew Effect', as it is written in the gospel according to Saint Matthew "For whosoever hath, to him shall be given, and he shall have more abundance: but whosoever hath not, from him shall be taken away even that he hath".

Differential teacher behaviours will sometimes result in differential pupil behaviours which may produce differential levels of achievement. Davidson and Lang (1960) have shown that children's self-concepts, classroom behaviour and academic achievements are related to their perceptions of their teachers' approval or disapproval of them. Other studies also indicate that if a teacher undervalues a child and unambiguously conveys his low expectations of him, the child will develop a poor concept of himself (Nash, 1973; Palfrey, 1973). Poor self image can result in low achievement motivation and hence low levels of achievement.

Turning to the link teacher expectation-teacher judgement, Cuttance (1980) states quite emphatically that "Teacher expectations influence their evaluations of pupils performance." However, research in support of this link is far from plentiful. Furthermore, it refers to expectations arising from a very limited range of information sources. There is some evidence that teachers discriminate among pupils on the basis of their social class. Work by Goodacre (1968) showed that the reading performance of children who were perceived to be working-class was assessed less favourably than that of middle-class children. However Williams (1976) working with data from 10,500 Canadian secondary pupils, concluded that teachers' assessments of their pupils' performances were not directly influenced by knowledge of the pupils' backgrounds. The results indicated that the teachers formed expectations based upon achieved pupil characteristics, such as IQ, past performance, ability grouping and educational ambitions. These expectations affected the teachers' evaluations of their pupils' current performances, as signified by course grades awarded. However, teacher expectations for pupil performance had only minor effects upon student learning, as measured by standardized achievement tests. Williams concluded that "Teacher expectations affect not so much what is learned in school as the certification of this learning by the school."

Perceived pupil ability has been shown to influence teachers' judgements in experimental conditions (Babad, 1980; Hughes et al., 1983; Smith, 1980). Chase (1979) reported that essays supposedly written by high achieving students received significantly higher marks than the same essays supposedly written by low achieving students. Cahen (1966) also found that the scoring of tests from hypothetical students could be biased by supplying information regarding the supposed ability of the students. Furthermore, the greater the amount of information given, the more likely were the marks to be biased in a direction consistent with the information given. Schrank (1968, 1970) obtained similar results in a natural experiment. US Air Force Academy instructors were led to believe that randomly grouped enlisted airmen were grouped according to ability. In this case, the groups labelled 'higher ability' received higher average grades in mathematics. But when the instructors were correctly informed that the men were not grouped according to ability, then all the groups obtained comparable average grades. These findings clearly show how teachers can mediate expectation effects by way of their grading practices, for in both experiments the students believed that they were grouped according to ability. Further evidence that teachers' expectations can influence their judgements appears in section 2.5.2.2.1. These studies also indicate that teachers' perceptions and expectations of children do not arise as the result of single variables in isolation, but that they are probably the result of complex interactions between pupil variables.

The link teacher judgement-pupil achievement needs little discussion. The possibility obviously exists that the performance of some pupils could be inaccurately assessed as a result of the teacher's judgement being influenced by expectancy effects. Moreover, inaccurate measures of achievement could lead to a pupil being placed in an inappropriate stream or set, which in turn could initiate further self-fulfilling prophecies.

Some workers cast serious doubt upon the logical and empirical basis of most of the expectancy literature. They question the causal relationship between pupil achievement and teacher expectation, and argue, with considerable justification, that pupil achievements are more likely to determine teacher expectations than the reverse causal sequence (West & Anderson, 1976). In an attempt to establish the direction of this causal relationship, Humphreys and Stubbs (1977) used a cross-lagged panel correlation method to reanalyse data relating to American high school students. This statistical technique allows the preponderant direction of causality between two correlated variables, which have been measured at two or more intervals of time to be inferred. The cross-lagged correlations showed that student achievement (grade average) caused various expectation measures - teacher expectation, student expectation, school expectation. Crano and Mellon (1978) employed the same statistical technique to analyse Barker Lunn's (1970) data on British primary children. In this study the results indicated that "The preponderant cause in the achievement-expectancy relationship was that of teachers' expectations causing children's achievements to an extent appreciably exceeding that to which children's performance impinged on teachers' attitudes". The conflicting conclusions from these two studies could reflect differences between the educational systems of the two countries, differences between primary and secondary schools (for example size, quality of teacher-pupil interaction), differences between the two student populations (for example age, academic ability, socio-economic status) or differences between the two teacher samples (for example educational level, teaching experience). More such studies are needed in order to clarify the direction of the causal relationship between teacher expectation and pupil achievement, and also to investigate whether certain variables, such as those suggested above, do in fact influence this causal interplay.

Although causal explanations of the expectancy-achievement relationship have not yet been fully investigated and are still being hotly debated, the existence of a relationship between teacher expectations and pupil achievement is well established.

2.4.3 Summary

1. Teacher expectation and the self-fulfilling prophecy are two of the terms used to describe the capacity of teachers to create a reality that corresponds to their perceptions.

2. The impact of interpersonal expectations has been demonstrated in both psychological and educational research. Evidence from educational studies is somewhat inconclusive, but a number of studies have shown that teachers' expectations can affect pupils' achievements.

3. A model is proposed in which information about a pupil generates expectations on the part of the teacher for that pupil, which in turn influence the achievement level of the pupil. ~~and subsequently his/her intelligence~~. Expectation effects are mediated through biased judgements of the pupil's behaviour and performance, and/or differential teacher behaviour. Evidence is cited in support of the various links making up this model.

4. Some researchers question the causal relationship between teacher expectation and pupil achievement. They argue that pupil achievements are more likely to determine teacher expectations than the reverse causal sequence. Attempts to resolve the direction have been inconclusive. However, the existence of a relationship between

teacher expectation and pupil achievement is well established.

5. Very few studies have investigated the expectations of science teachers, and whether they are similar for both male and female pupils.

2.5 SEX BIASED JUDGEMENTS

This section is divided into two parts - one dealing with sex bias in the evaluation of performance in a variety of situations, the other dealing with sex bias in teachers' marks. It is important first to demonstrate that sex biased judgements can be made by many types of people in all sorts of situations. If these findings are taken to indicate the way in which our society behaves, then since teachers are members of this society, they are likely to behave similarly. The second part looks specifically at evidence of teachers' sex biased judgements in their working lives.

2.5.1 Differential evaluation of males and females

Recent research in the area of perceived sex differences suggests that the performance of males and females may be differently perceived and thus differently evaluated. Biased evaluations encompass all aspects of performance, including any resulting products. Most studies indicate that both males and females give similar evaluations. Thus it seems that the sexes share the same biases.

2.5.1.1 Performance

Several studies have indicated that the performance of a male tends to be rated more favourably than the same performance by a female. Besides viewing male performance as better, the task and the performer are also rated higher when the task is performed by a male (Feather & Simon, 1976; Deaux & Emswiller, 1974). Successful performance by a male on a masculine task will probably be attributed to skill, whereas the same performance by a female is more likely to be explained by luck (Deaux, 1976; Deaux & Emswiller, 1974). See also section 2.3.1. The reverse is not true for performance on a feminine task. However, the performance of a woman may be viewed as reflecting more effort (Feldman-Summers & Kiesler, 1974; Taynor & Deaux, 1973).

Taynor and Deaux (1975) suggest that studies of perceived sex differences in performance should consider a range of influential variables, such as the sex stereotype of the task, the sex of the person who performs the task, and the manner in which the task is performed. Their study indicated that a feminine task is undervalued in various ways.

Although a general devaluation of female performance in relation to male performance is conveyed by the literature, some studies suggest that bias can work both ways (Deaux & Taynor, 1973). Feather and Simon (1975) and Feather (1978a) found a tendency for successful females to be downgraded in relation to successful males, but for unsuccessful males to be downgraded in relation to unsuccessful females.

Research into the evaluation of job performance has yielded conflicting results. Feldman-Summers and Kiesler (1974) found that male, but not female college students perceived a female doctor to be less able than a male doctor. In education, some studies have shown that male teachers are rated more favourably (Wilson, 1974), others have shown that female teachers are rated more favourably (Mackie, 1976) and yet others have shown no difference (Elmore & La Pointe, 1974; Hesselbart, 1977). More complex investigations have demonstrated the influence of student sex (Bray & Howard, 1980; Ferber & Huber, 1975; Kaschak, 1978), student age (Goebel & Cashen, 1979), teacher's subject (Ferber & Huber, 1975; Masters, 1978), and teacher's teaching style (Harris, 1975) upon evaluations of teachers. Detailed questionning about different aspects of teaching performance has helped to clarify the perceived strengths and weaknesses of male and female teachers. Male teachers are often judged to be more competent with respect to the academic and pedagogic facets of the teacher's role (Stanworth, 1981), but female teachers are usually rated higher on interpersonal dimensions, especially warmth (Bennett, 1982). Turning to the job of library administrator, an occupation that is generally viewed as being sex neutral, Brief and Wallace (1976) found no difference in the evaluation of male and female performances. However, in the female dominated profession of nursing, female nurses are rated as being more competent than male nurses (Winkler, 1982). The performance of women working as grocery clerks, a low status job, is also rated higher than that of men (Hamner et al., 1974). In conclusion, most studies indicate that behaviour conforming to sex role prescriptions is viewed more favourably than behaviour more typical of the opposite sex (Feather, 1978b; Feather & Simon, 1975; Stiver, 1976), but there are exceptions (Taynor & Deaux, 1973, 1975).

2.5.1.2 Products

In 1968 Goldberg published a provocative study which showed clearly that male output is evaluated more favourably than is identical female output, both in relation to the task and the performer. The study explored the prejudice of women against other women in areas of intellectual and professional competence. Goldberg hypothesized that women would value the professional work of men more highly than the identical work of women. Female college students were asked to evaluate six abridged journal articles. Two of the articles were on law and city planning, fields previously judged to be masculine; two were on education and dietetics, fields previously judged to be feminine; and two were in neutral fields - linguistics and art history. Three of the articles bore male names and three bore female names, but only half of the students saw a particular article ascribed to a male author, whereas the other half saw it ascribed to a female. The articles supposedly written by males were judged to be of greater value, and the authors more competent, than those articles supposedly written by females. This biased judgement arising from distorted perceptions constitutes a prejudice - anti-feminism. Contrary to a second hypothesis of Goldberg, that the tendency of women to devalue the products of women would diminish or be reversed in traditionally feminine fields, it was found that prejudice against women extended even to elementary school teaching and dietetics.

Goldberg's findings have tended to be viewed as definitive by many people, but some workers have questioned the statistical significance of the results (Mischel, 1974), some have doubted the generality of the conclusions (Cline et al., 1977; Pheterson et al., 1971), and others have questioned the notion of an all pervasive devaluation of women regardless of topic and regardless of questions asked (Ward, 1981a). Such queries have resulted in further investigations into the effect of various factors upon the biases reported by Goldberg, in an attempt to repeat and extend his conclusions. Probably the major contribution of Goldberg's study has been to stimulate further work into perceived sex differences.

Using Goldberg's experimental design, Mischel (1974) increased the number of variables in order to investigate whether the subject's sex, educational level and cultural background affect the tendency to evaluate articles on the basis of the author's sex. The findings indicate the existence of much greater specificity in professional evaluation than the diffuse, non-specific upgrading of male competence reported by Goldberg. Mischel found evidence of sex bias in the evaluation of the journal articles, but the subjects tended to prefer articles ascribed to male authors in the male fields and articles ascribed to female authors in the female fields. This bias was shared by both male and female raters. In contrast, the educational level of the rater (high school or college) did affect the evaluation of the articles, with specific and complex interactions appearing. The cross-cultural data, comparing American and Israeli students, indicated that the cultural background of the rater also appears to affect ratings of competence, but not views on the sex typing of occupations. A recent study by Gross and Geffner (1980) confirms Mischel's findings that sex role prejudice is a complex phenomenon involving many factors.

Concentrating upon age as a variable, Etaugh and Rose (1975) investigated whether adolescents (13, 15, 17 years old) would differentially evaluate magazine articles. The entire age range and both boys and girls displayed sex bias in their evaluations. This sex bias operated in both directions, but most frequently the work of male authors was preferred. This upgrading of articles by male authors was most pronounced in masculine fields. Like adults, adolescents apparently devalue the products of females more frequently and in more ways than the products of males.

Gold (1972) working with American university students found that not all of the female sample devalued female authored articles, and that the male students actually viewed female authors more favourably than male authors. In contrast, Ward (1981a) reported that male students at a British university rated female authors significantly lower on status and competence, but neither sex judged the quality of the essays written by females less favourably.

Some studies have failed completely to demonstrate any form of differential evaluation (Baruch, 1972; Dansker, 1974; Levenson et al., 1975; Wittekind, 1975). When Pheterson (quoted in Pheterson et al., 1971) asked uneducated, middle-aged women to rate articles on marriage, child discipline and special education, their evaluations were not sex biased. To explain this result, Pheterson suggested that the women probably regarded the very publication of an article as a considerable accomplishment. This would imply that female achievements are evaluated similarly to male achievements once they have already been proved successful. To test this hypothesis, Pheterson et al. (1971) asked college women to evaluate paintings which were either presented as entries in an art competition or as award winning entries. The award winning entries of females were rated as highly as similar entries by males, but the other entries of females were devalued in comparison with the male entries. These results support Pheterson's proposal that male achievements are judged more favourably simply because males usually have greater chances of succeeding, but that once females are seen to be successful, then their achievements are judged fairly.

Etaugh and Sanders (1974) were unable to repeat Pheterson et al's (1971) findings. However, although their study showed no evidence of sex bias in ratings of quality, creativity or artistic future, they did find that sex typed characteristics were differentially evaluated. Subjects applied sex role stereotypes in judging the competence and emotionality of opposite sex, but not same sex, artists. Using a similar experimental design, Ward (1979a) also failed to reproduce previous findings. She found that art students rated more favourably a high quality painting attributed to a male artist and an inferior painting attributed to a female artist. The reverse was true for university students.

These numerous studies employing the same basic experimental design as that adopted by Goldberg have produced diverse and often conflicting results. Many factors could account for these differences. Ward (1979b) suggests that rater sex, rater expertise, sex appropriateness of the achievement, level of the achievement, and ambiguity concerning the qualifications or status of the person producing the work, all influence the way in which the achievements of males and females are evaluated. Nevertheless, regardless of such specific interactions between variables, most studies which have employed Goldberg's experimental design, have reported sex biased evaluations of some description. Furthermore, there are more reports of women's achievements being devalued in comparison with men's achievements than the reverse.

2.5.2 Sex Bias in Teachers' Marks

There is considerable evidence indicating that teachers, as a group, are no more exempt from bias in their judgements than are other people. Data were originally gathered from natural school settings, the behaviour of teachers in their classrooms was observed, and the marks that teachers had awarded their pupils were inspected. More recently teacher-pupil interactions have been recorded more objectively using observation schedules, and the marks that teachers award have been investigated using controlled experimental techniques. Data from both natural and experimental sources reveal that teachers often respond to a variety of extraneous factors when assessing their pupils. The extraneous factors may cause many teachers to perceive differences in their pupils' potentialities and competencies, with the result that the extraneous factors often influence the teachers' assessment of their pupils' behaviour and academic achievements.

A study by Carter in 1953 revealed that non-cognitive variables, such as pupil sex, socio-economic status and personality, contributed to the assignment of marks by teachers of elementary algebra. A year later, Hadley (1954) was writing that:

> Among the many factors reported in the literature to be included in the marks teachers assign their pupils are - actual attainment, the teacher-pupil relationship, deportment, sex, promptness and attendance, personal appearance, obedience, effort and attitude.

Teachers' biased judgements are not just of academic interest but rather a major educational concern, for they may directly affect the self concepts and academic careers of many schoolchildren. Thus over the years many studies have attempted to investigate the validity of the marks that teachers assign to written work, and to ascertain the contribution of non-cognitive factors, such as pupil sex, to those marks.

2.5.2.1 Evidence from School Records

Regular investigations into sex bias, one of the more psychologically plausible and ostensibly one of the more readily investigable biases, have been recorded over the last fifty years. Most of these studies have indicated that pupil sex can influence teachers' evaluations of written work in natural school settings. Although some of the studies were conducted many years ago and most of them refer to American grading practices, a selection will be briefly reviewed in order to indicate the range of questions that workers investigated, the procedures employed and the findings reported.

Teachers generally award higher marks to girls than to boys. However, it is questionable whether these differences in evaluation reflect substantial differences in performance. Garai and Scheinfeld (1968) quote three studies from the fifties and sixties which showed that boys obtained higher scores on standardised achievement tests than teachers awarded them in class. Carter (1952) reviewed eight studies from the thirties and forties which had investigated grading practices in high schools. Seven of the studies revealed that girls received higher average marks than boys, and five of them indicated that this was a reflection of sex bias amongst the teachers. Female teachers displayed the greater bias by both overrating girls and underrating boys.

Most of the early studies did not take the intelligence of the pupils or their level of attainment on standardized tests into consideration, an omission which must detract from their conclusions. The work of Carter (1952) does consider intelligence and achievement test scores as variables, and therefore the findings have greater validity. He investigated whether teachers tend to favour one sex and whether the teacher's sex determines the pupil sex preferred. Comparisons were made of teachers' marks, achievement test scores and IQ scores for nine elementary algebra classes, four of which were taught by women and five by men. Although there were no significant differences in intelligence or algebra achievement score, significant differences were found in the marks assigned by the teachers. The girls were given higher marks than the boys, and the female teachers tended to award higher marks than the male teachers. Thus the pupil's sex did influence the assignment of marks, but there was no interaction between pupil sex and teacher sex. In addition, the data showed that the teacher's marks reflected not only algebraic achievement but also intelligence and other factors, of which only pupil sex was determined.

More recent studies have confirmed that boys are more likely to underachieve at the secondary level than are girls. Sexton (1969) found that more boys than girls were being awarded lower ratings by their teachers than their scores on aptitude tests would indicate. McCandless et al. (1972), whilst investigating the effect of pupil sex upon teachers' marks, included two more variables - race and social class. Data on intelligence, standardized achievement and teachers' academic subject marks were collected for 7th grade pupils (12 years old). It was found that the boys scored slightly lower marks than the girls on the standardized tests, but that they received much lower marks than the girls from their teachers. Analysis revealed that only 4% of the variance in the teachers' marks assigned to boys could be accounted for by their standardized achievement test scores, whereas about 15% of the variance for girls could be explained. However, these figures conceal great differences between the different sub-groups. Correlation coefficients for achievement test scores and teachers' marks ranged from -0.17 for advantaged white boys to 0.73 for disadvantaged black girls. Although teachers' marks did not correlate consistently with achievement test scores, they were positively correlated, at a modest level, with intelligence. In spite of complex interrelationships between the different variables, the results confirmed previous findings concerning the biasing influence of sex upon teachers' marks, and also suggest that advantaged children are assigned marks according to their intelligence and socialization patterns, whereas disadvantaged children are assigned marks according to their intelligence and level of school achievement.

At the primary level, Arnold (1968) investigated the effects of pupil sex and social class upon the marks awarded to 10 and 11 year old pupils. For each of four academic subjects (reading, spelling, language and arithmetic), pupils were divided into three categories on the basis of achievement test scores. Teachers' marks for each subject were then collected from report cards and subjected to analysis of variance. It was found that social class did not significantly affect the teachers' marks, but pupil sex did. In all four subjects, the mean school marks assigned to girls were higher than those assigned to boys. An earlier study by Hadley (1954) of female teachers, produced similar effects for pupil sex. Teachers' term marks and the achievement test scores of 4th, 5th and 6th graders in the four basic skill areas of reading, arithmetic, language and spelling were investigated. 45% of the girls studied received teachers marks higher than their achievement test scores warranted, and 40% of the boys received teacher marks lower than indicated by their achievement test results. Overall correlations between teachers' marks and objectively measured attainment in the twenty classrooms studied ranged from 0.20 to 0.94.

The favouring of girls in the allocation of grades occurs at all educational levels and in a wide range of subjects, including those in which males usually excel (Maccoby, 1967). Consequently, in the American educational system, girls receive higher grade composites or grade-point averages than do boys of similar intelligence (Dwyer, 1979). The gap between the grade-point averages of girls and boys is greatest at the primary level and it steadily narrows throughout secondary school (Garai & Scheinfeld, 1968).

A recent British study by Bradley (reported by Tysoe, 1982) into marking practices at university level revealed a bias in favour of male students. Bradley compared the marks awarded to the projects of final year undergraduates by external examiners and project supervisors. When the projects bore the students' names, the external examiner tended to give lower marks to the competent female students than the project supervisor. In contrast, there was no evidence of sex biased marking by the external examiner when the scripts carried no names.

2.5.2.1.1 Explanations

Several explanations have been put forward to account for girls' greater success in exercises and examinations which are set and marked by their teachers. It has been suggested that female teachers discriminate against boys, and since the majority of primary teachers are female, their biased treatment of boys could explain the significantly poorer marks received by boys at the primary level (Sexton, 1969). Several studies have shown that female teachers are less favourably disposed towards boys and male qualities than towards girls and female qualities (Datta et al., 1968; Hart & Olander; 1924; Jackson et al., 1969). Moreover, teachers are more likely to underrate the intelligence and potential of boys than of girls (Doyle et al., 1972).

The value of the female teacher bias explanation can be assessed by considering studies that have compared the marks awarded by both female and male teachers, since the female teacher bias explanation implies that male teachers are much less discriminatory against boys. Evidence concerning the comparability of the marks assigned by male and female teachers to male and female pupils is rather confusing. Some studies have found that male teachers generally gave higher marks than female teachers (Arnold, 1968; Douglass & Olson, 1937), but other studies have found that female teachers gave higher marks to both girls and boys than did male teachers (Carter, 1952; Newton, 1942). The findings of a few studies have suggested an interaction between teacher sex and pupil sex. Edmiston (1943) found that boys received higher grades from male teachers than from female teachers, but that girls received lower marks from male teachers. In summary, very few studies have produced evidence of an interaction between teacher sex and pupil sex in the allocation of marks, and the question of whether female or male teachers tend to give higher marks is unresolved. However, all studies have found that the work of girls is assessed more favourably than that of boys by both female and male teachers. Thus, if teachers do discriminate against the work of boys, it is not just female teachers who discriminate, but both sexes.

A second explanation for girls' higher marks from teachers is that girls over-achieve in comparison to boys of similar ability. Stanley (1967) proposes that girls are "more conscientious students than boys, working more nearly in accordance with their basic abilities and skills". Not only are girls more conscientious about the completion of set work, but they also conform better than boys to classroom rules and norms. The compliant and cooperative behaviour of girls may predispose teachers to also view their work favourably. Thus a third explanation for girls' higher marks is that teachers are influenced by pupil conduct when assigning marks in academic subjects (Garai & Scheinfeld, 1968). In support of this explanation, Brophy and Good (1974) report work showing that the conduct grades of boys were more highly correlated with academic subject grades than was the case for girls.

A fourth explanation focuses upon the finding by several researchers that boys are more successful than girls in answering multiple choice questions, whereas girls appear to excel on essay type questions (Harding, 1979; Hoste, 1982; Murphy, 1982). Assuming that the aptitude test scores in the studies described above were obtained almost exclusively from objective (multiple-choice) tests, and such is very likely, then boys would be at an advantage and would be expected to achieve higher scores than girls. On the other hand, if the work and examinations set by teachers is almost exclusively of a non-objective nature, then girls would be at an advantage and would be expected to achieve higher marks than boys. The appeal of this explanation is that it neatly accounts for both girls' higher marks from teachers and boys' higher scores on aptitude tests, without accusing teachers of discriminating against boys or being unduly influenced by a pupil's sex or conduct when assigning marks.

2.5.2.2 Evidence from Controlled Experiments

The evidence from school records reviewed in the above section indicates that boys consistently receive lower marks than girls even when equal in intelligence quotients and in achievement test scores. This prejudice against boys on the part of school teachers is commonly attributed to the single factor of sex. However, as suggested in the previous section, the variable 'sex' is probably inextricably linked with a large cluster of other closely interrelated variables, such as conduct, effort, attitude to work and school, rapport with teachers, co-operativeness, obedience, motivation, appearance. In natural school settings it is impossible adequately to separate out all these interrelated variables. To overcome this problem, many researchers have employed experimental designs which incorporate adequate controls.

2.5.2.2.1 Single Factors

Controlled experiments which effectively separate identifiable factors have shown that a single factor is capable of influencing teachers' judgements to the extent that otherwise identical samples of work will be marked differently. Few workers have studied pupil sex as a single factor, but it has been included as a factor in more complex experimental designs which investigate interactions between several factors (see section 2.5.2.2.2).

Spender (1982) manipulated pupil sex by asking teachers to mark work (essays, projects, assignments) that was presented to some as being the work of a boy and to others as being the work of a girl. She repeated the experiment on five different occasions and each time found that the teachers rated work attributed to boys more highly than work attributed to girls. Spear (1982, 1984) conducted a similar experiment with science teachers as markers. They were asked to rate six experimental write-ups on a number of work characteristics. However, pupil sex was varied so that each work sample was presented to half the teachers as being the work of a girl and to the remaining teachers as being the work of a boy. It was found that work attributed to a boy was generally rated higher for scientific accuracy, richness of ideas and organisation of ideas than identical work attributed to a girl. The only work characteristic on which girls were favoured was neatness.

The remaining single factor experiments reviewed below, investigate factors other than pupil sex. They are included in this survey since they illustrate the range of experimental designs which have been exployed. In addition, the results are of considerable value and importance to any investigation into teachers' marks, and are of particular relevance to the design of the study which follows.

Investigations into interpersonal relationships have indicated that names are associated with stereotypes (Buchanan & Bruning, 1971; Bruning & Albott, 1974; Duffy & Ridinger, 1981; Garwood et al., 1981), and that these stereotypes can influence our perception of a person (Busse & Seraydarian, 1979; Lawson, 1971; Leirer et al., 1982; McDavid & Harari, 1966). The influence of name stereotypes appears to operate in educational, as well as social situations. Nelson (1977) reported work which suggests that name stereotypes might influence the standard of academic achievement teachers expect of their pupils. Harari and McDavid (1973) investigated whether such stereotyped expectations were also reflected in teachers' assessments of student work. Experienced female primary teachers were asked to mark eight essays of similar standard, which were ascribed to boys and girls with names which had previously been independently judged as desirable or undesirable. The essays ascribed to children with the popular names received significantly higher marks than those supposedly written by children with the unpopular names, the difference being greater for the boys' names than the girls' names.

Scannell and Marshall (1966) investigated the effect of errors upon the grade awarded to an essay. The errors selected were those of grammar, punctuation and spelling. The judges were instructed to mark for content only, but it was found that the copies of the essay with errors received lower grades than the perfect copies. Subsequently, Marshall (1967) found an inverse relationship between the grade given to an essay and the number of errors it contained.

Several writers have investigated the combined effect of writing errors and various other factors upon essay scores. Freedman (1979) found that punctuation and spelling errors, in combination with other composition variables, produced main effects upon ratings of the essay, and also interaction effects. Marshall and Powers (1969) combined different types of linguistic errors with different standards of handwriting neatness. Significant differences were found between the mean grades awarded to an essay written in neat handwriting and an essay written in mediocre handwriting, and between an essay containing no linguistic errors and an essay containing spelling and grammatical errors. The findings of a study by Chase (1968) provide additional evidence of the biasing effect of handwriting quality, but not of spelling accuracy. Chase also detected interaction between quality of writing and the order in which essays were presented. A further investigation by Marshall (1972) into the effects of writing neatness and number of errors, failed to produce any significant differences or interactions between the experimental variables.

Studies which have investigated handwriting quality as a single factor have generally produced more unequivocal results. Markham (1976) showed that essays written in better handwriting styles consistently received higher marks than essays written in poor handwriting styles, regardless of the quality of the content. Furthermore, teacher characteristics such as age, degrees held, teaching experience, level taught (1st-5th grade) did not significantly influence the marks assigned to the essays. Briggs (1970, 1971) showed that handwriting style can act as a biasing influence upon British teachers' assessments of essays. Panels of primary and secondary English teachers were asked to impression mark ten essays, each of which differed in content and handwriting style. Analysis of the results revealed that essays written in preferred handwriting styles received higher marks than those written in lower ranking styles, and the variation in marks increased as the quality of the essay decreased. One essay received a mean mark of 13/20 when written in the most popular writing style, but only 10/20 when written in the least popular writing style. More recently Briggs (1980) reported that the effect of handwriting style may be so influential as to determine whether a candidate passes or fails in an external examination.

Soloff (1973) employed a vary simple experimental design to study the general effect of tidiness upon teachers' assessments. Essays which had been produced in two versions - neat and 'sloppy' (words crossed out) - were graded for content. The mean grades for the neat versions were significantly higher than for the sloppy versions. Soloff concluded that factors other than content, namely handwriting and the general appearance of the essays, had influenced the teachers in their assessment. Similar findings and conclusions also resulted from a British investigation in which science teachers were asked to mark experimental write-ups (Rivett, 1980). Although the visual presentation of a piece of work clearly influences the mean mark awarded to it, the effect may not be as universal as the above results suggest. Huck and Bounds (1972) found that markers who were themselves neat writers tended to penalize untidy work, whilst markers whose own writing was untidy were not affected by the neatness of the work they were marking.

Spender (1977) besides investigating the influence of neatness upon teachers' marks, also included the variable pupil sex. She found that work attributed to a boy received a higher grade regardless of whether it was tidy or untidy. The findings of this unpublished study are more noteworthy for the sex bias that is reported than for neatness induced biases.

In summary, research has shown that teachers can be influenced in their judgements of written work by a variety of non-cognitive variables relating to both the characteristics of the script and the author. Author characteristics that have been shown to bias teachers' evaluations include students' first names, their physical attractiveness (Holahan & Stephan, 1981; Landy & Sigall, 1974), their sex, and their race (Guttmann & Bar-Tal, 1982). An even greater number of extraneous variables associated with the script can influence teachers' judgements. They include handwriting quality, standard of presentation, linguistic errors, readability level (Chase, 1983), length of essay (James, 1976), essay title (Wiseman & Wrigley, 1958), and order of marking (Daly & Dickson-Markman, 1982; Hales & Tokar, 1975; Hughes et al., 1980).

2.5.2.2.2 Multiple Factors

Several studies have indicated that teachers' perceptions and evaluations of pupils' achievements do not arise solely as the result of single extraneous variables operating in isolation, but that they are more likely to be the result of complex interactions between pupil variables.

In this country, Bull and Stevens (1976, 1979) have investigated the interaction between the biasing influences of handwriting style and a writer's sex and physical appearance. They asked teachers and students to grade an essay, which was accompanied by a report card containing a photograph. Both the essay and the report card contents were identical except for the following variables - sex of writer, attractiveness of writer, and penmanship. After analysing data for the male and the female writers separately, no effects of attractiveness or penmanship were found when the writer was supposedly male, but both penmanship and attractiveness influenced the ratings given to the female authored essay. Moreover, the teachers were influenced more than the students by the handwriting and attractiveness of the writer. Unfortunately, the design of the experiment and the small sample used restrict the generality of the results. "Nonetheless, the study draws attention to the complex relationship which exists between physical attractiveness, presentation and the sex of the writer and the expectations about essay content that these external features can set up in an examiner" (Wade, 1978). Although such relationships give cause for concern, they are not new findings. Back in 1947, Ross complained "It seems too bad that the marks received by certain individuals are conditioned more by the contours of the face than by the contents of the head".

Finn (1972) attempted to gain insight into the manner in which the pupil variables of sex, race and ability interact with the environmental variable of school locale to influence teachers' expectations regarding their pupils' achievements. Teacher expectation was judged from teacher ratings of sample essays. Experienced 5th grade teachers from urban and suburban schools were asked to rate two short essays, each of which was accompanied by a covering letter providing information about the author's race (Negro-white), sex (male-female) and ability (high-low). The results were subjected to non-orthogonal multivariate analysis of variance. This showed that the main effect of pupil ability was not significant. However, pupil ability did interact with teacher locale. The urban, but not the suburban teachers, tended to rate essays ascribed to high ability pupils more favourably than identical essays ascribed to low ability pupils. Similarly, although pupil sex and race did not significantly affect teacher expectancy when considered alone, there was a significant interaction between these variables and teacher locale. The urban teachers gave significantly higher ratings to white male than to white female authors, but were not biased in their judgement of negro males and females. The suburban teachers were generally unbiased in their judgements, although there was a slight tendency to award higher ratings to negro pupils than white pupils. Finn concluded that:

> In certain settings teachers do hold differential expectations for the achievement of student groups having common non-achievement characteristics. In a setting characterized by relatively low teacher expectations in general, expectations for specific pupils were so pervasive as to bias evaluations of pupils, even where the achievement evaluated was identical from one individual to another. (p. 407)

Kehle et al. (1974) carried out a similar experiment with the objective of determining the way in which pupil sex, race, intelligence and attractiveness interact to influence teacher expectations. Urban primary teachers were supplied with a colour photograph of a pupil which conveyed information concerning sex, race and physical attractiveness, and a psychological report which included an IQ score. The teachers were then asked to assess the pupil's personality characteristics, and to judge the pupil's academic performance by rating an essay. Of the four pupil variables under investigation, only sex produced a main effect in the ensuing multivariate analysis. However this significant effect was only upon the personality rating, not upon the essay rating. The significant effects which were found for the sex by attractiveness interaction did refer to essay assessment. Attractive females were rated more highly on the essay than unattractive females, whereas the reverse was the case for males. Teacher judgement was also influenced by the four-way interaction, personality ratings were significantly affected and the effect on essay ratings was nearly significant. Essays supposedly written by white males tended to be rated lower than ones attributed to white females, with the exception of unattractive pupils of low intelligence. In contrast, both male and female negro ratings were equivalent, except for unattractive pupils of high intelligence. These findings, together with those of Finn's study, indicate that non-cognitive variables, which individually have been implicated as contributing towards teacher expectation and teacher bias, most probably combine and interact in a complex and not readily predictable fashion in natural classroom settings.

The above two studies by Finn (1972) and Kehle et al. (1974) provide valuable data and results concerning the biasing influence that a pupil's characteristics can exert upon a teacher's assessment of that pupil's written work. However, the assumption in both studies that the experimental design was measuring teacher expectation, rather than teacher bias, appears somewhat tenuous. Williams (1976) has suggested that teacher expectation may be defined as "expectations for the student's future academic achievement". This particular variable was not measured directly in either of the experiments. Instead, Finn and Kehle et al. assume that by investigating teachers' judgements they are indirectly measuring their expectations. But expectations do not of necessity lead to biased judgements and even when this is the case the exact relationship is often elusive. As Kehle et al. write:

> A low score on the essay can reflect either positive or negative teacher expectation. The low essay score can be congruent with bias. A low essay score can also reflect the teacher's disappointment or surprise with the quality of the essay given a positive expectation for student performance. (p. 59)

In view of the difficulties associated with attempting to relate observed effects and teacher expectation, it might have been more meaningful if Finn and Kehle et al. had simply discussed their results in terms of teacher bias rather than teacher expectation.

2.5.3 Summary

1. A man's performance at a task tends to be rated more favourably than the same performance by a woman. A number of variables influence the way that male and female performances are evaluated, including the sex stereotype of the task, the sex of the person performing the task, and the manner in which the task is performed.

2. The achievements of a man (i.e. journal articles, magazine articles, paintings) are often judged to be better than identical achievements from a woman. Moreover, the male author/artist is also judged to be more competent. The upgrading of articles by male authors tends to be most pronounced in masculine fields. A number of variables influence the way in which the achievements of men and women are evaluated, including the sex of the rater, the expertise of the rater, the sex appropriateness of the achievement, the level of the achievement, and any ambiguity concerning the qualifications or status of the person who produced the work.

3. A number of studies, mostly American, show that girls receive

higher average marks from their teachers than do boys at both the primary and the secondary level. If the marks that pupils receive from their teachers and the scores that they obtain on aptitude tests are compared, then the indications are that teachers overmark more girls than boys.

Four explanations to account for this phenomenon are advanced:

(a) Female teachers discriminate against boys.

(b) Girls over-achieve.

(c) Teachers are influenced by girls' superior conduct.

(d) Boys are better at multiple choice questions, whereas girls excel on essay type questions.

The last explanation is considered to be the most satisfactory.

4. Experimental studies indicate that teachers' assessments of written work can be influenced by a variety of non-cognitive variables relating to both the characteristics of the script and the author, e.g. the pupil's first name, his/her sex, handwriting quality, standard of presentation, length of essay. When pupil sex is manipulated, work attributed to a boy is rated more highly than identical work attributed to a girl. There is evidence that science teachers conform to this marking pattern, but more needs to be known about the sex biased marking patterns of science teachers.

When several non-cognitive variables are investigated together, their combined effect upon teachers' marks is not always predictable, because of complex interactions. No studies have investigated the effect of interactions between pupil sex and other variables upon science teachers' marks.

CHAPTER 3

DEVELOPMENT OF RESEARCH HYPOTHESES

3.0 Contents

3.1 EVIDENCE FOR THE PROPOSED MODEL

The foregoing literature review clearly indicates that teachers do hold sex stereotyped beliefs, may hold different expectations for certain groups of pupils, and sometimes make biased judgements when marking. However, much of the evidence originates from the United States of America, and very few studies have involved science teachers. The next two sections will bring together the available evidence relating to science teachers. Both sections aim, not only to draw attention to evidence that is directly relevant to the separate topics identified in the proposed model (Figure 1.1), but also to refer to evidence that interconnects the separate topics by spanning the rather arbitary boundaries that delimited the topics in the previous chapter.

3.1.1 <u>Empirical evidence</u>

Several of the topics identified in the model (Figure 1.1) have either not been researched, or else the research is very scant and incomplete. Science teachers' views about the gender connotations of the subjects that they teach have not been studied. Likewise, the attributions used by science teachers to explain the success and failure of their pupils have not been studied. A few workers have mentioned that science teachers hold relatively traditional sex stereotyped beliefs (Seale et al.,1982; Whyte, 1983a), but detailed evidence is lacking.

Evidence that teachers hold different expectations for the achievements of boys and girls studying science is provided by the results of a marking experiment (Spear, 1984). In the course of assessing samples of pupils' work, secondary school science teachers expressed higher expectations for 'boy' authors than for 'girl' authors, as signified by their judgements of the pupils' potential for O level physical science courses. Smail & Whyte (1983) examined sex bias in science teachers' expectations using a case study approach. They produced a written description of the qualities, aspirations and predicted examination results of a fifth-form pupil. This description was given to teachers who were assembled together for a workshop on Girls and Science. All the teachers received identical pupil descriptions, but half of the teachers were told that the pupil was a boy called Denis Johnson, whilst the other half were told that the pupil was a girl called Denise Johnson. The teachers were asked to record the career and/or educational advice they would give to the pupil. The replies provided evidence of sex typed expectations of boys and girls, with further education being deemed more important for a boy.

Evidence that science teachers' expectations actually influence their pupils' achievements is equally sparse. McDuffie (1979) found a significant correlation between what teachers expected pupils to achieve and pupil performance on a standardized science test. A more pertinent study has been reported by Rowell (1971). Although it was only a small pilot study involving twelve teachers, the findings do suggest that teachers' expectations can either reinforce or counteract girls' under-achievement in science.

There are few reported instances of pupil sex influencing the judgements made by science teachers. The work of Spear (1984), referred to previously (sections 1.2 and 2.5.2.2.1), suggests that teachers view the written work of boys more favourably and thus that boys' work may be awarded higher marks than similar work from girls. Pupil sex also influences the advice that pupils receive from teachers (Raat, 1983). Physics teachers demand higher marks from girls than from boys before recommending that they take physics.

Studies of science teachers' behaviour towards boys and girls are more plentiful. The recent use of classroom observation techniques has provided quantitative data showing that teachers interact differently with boys and girls in science lessons. Boys are more demanding of their teacher's attention than girls (Whyte, 1983b). They initiate more contacts with teachers and generally display more disruptive behaviour.

In response to such behaviour it is not surprising that teachers pay more attention to boys and have to reprimand them more.

Teachers also direct more teaching exchanges to the boys in their classes. Raat (1983) reports that more boys than girls answer questions posed in physics lessons. Imbalances between boys' and girls' participation in science classes are so common that teachers seem to accept them as natural. Teachers who do try consciously to involve girls more in their lessons often feel that they are being 'unfair' to the boys in the class (Whyte, 1983b).

3.1.2 Impressionistic evidence

Circumstantial evidence and anecdotal reports suggest that some science teachers hold greater expectations for the boys in their classes than for the girls. These different expectations sometimes lead to differential treatment of boys and girls.

Statements can be found in the literature which indicate that some science teachers do consider science to be a subject most suited to boys.

> Although physics and chemistry cannot really be thought of as boys' subjects any longer, a great many of the teachers still tend to regard them as such.
> Schoolgirl quoted in DES (1980, p.26)

> Male teachers in particular are inclined to make inadvertent comments in class which reflect an assumption that science is a boy's subject.
> National Council of Women of Great Britain (1982, p.6)

> Teachers frequently do not consider science to be all that important for a girl.
> Kelly (1979, p.109)

Moreover teachers, especially male teachers, tend to hold higher expectations for the boys in their classes than for the girls.

> Men say that they don't treat girls differently and believe that girls are equal with boys but why then when a female is top of his class in Physics do they fuss so much. They must have greater expectations of males than females.
> Female biology teacher quoted in Gannon (1980, p.27)

Several writers indicate that teachers often judge girls' attitudes and aptitudes for science to be inferior to those of boys.

> In his talk the head of science referred to girls being interested in biology - but when it came to physics and chemistry it was "The old, old story that they are not interested".
> ACE case study (1979, p.216)

> Many teachers are still imbued with the belief that although they are required to teach girls science, it is not surprising if the girls do not do well.
> Whyld (1980, p.8)

> ... - it was in Physics - and he (the physics teacher) said, "Now if you were boys you would understand this."
> Schoolgirl quoted in Stanworth (1981, p.37)

As a consequence, girls receive less encouragement than boys to continue their study of science.

> Teachers and parents frequently have stereotyped ideas of what is suitable and interesting for each sex, and so may not encourage girls to persevere with science.
> Kelly (1981, p.283)

> The boys were encouraged more and the girls were not actually put down but were, shall we say dampened.
> 16-year-old girl quoted in Kelly (1981, p.235)

> The teacher took no interest in the girls what so ever. He cared about the boys' future and not the girls.
> 16-year-old girl quoted in Kally (1981, p.238)

Some girls even encounter active discouragement from their science teachers.

> Many of the girls I spoke to reported discrimination from their physical science teachers who either ignore them or indicate that girls are not expected to understand the principles of science.
> Gannon (1980, p.27)

> I was recommended for 'O' level Physics along with many other girls in the class, but our teacher put us off, "You don't want to take Physics, it's a boys' subject," he said. "You'll find it very boring and difficult."
> 16-year-old girl quoted in Kelly (1981, p.240)

> Men teachers at my school seemed to put people off. They used to say to girls, "don't worry if you find physics difficult, there are more important things in life than physics."
> Female undergraduate quoted in Lewis (1983, p.191)

> Some masters have told me (at ASE meetings) that they deliberately frighten off the girls - especially in the first term of the sixth form - by making the work difficult. Girls who are more conscientious and less confident drop out, and then they revert to the normal standard of work.
> Physics teacher quoted in Kelly (1981, p.258)

In a few cases, teachers' feelings against girls in science are so strong that they lead to prejudice and hostility.

> It is a fact that most girls have not the type of mind that faces a problem, nor reasons well from given data - not even my star girls, who got as far as Oxford and/or Cambridge, and to a first and a PhD in one case. Even this girl just could not compare with her boy rivals but she took the subject further and at a university where standards were lower.
>
> Male physics teacher quoted in Kelly (1981, p.258)

> There is evidence of hostility towards girls in science from some men teachers.
>
> Kelly (1978a, p.66)

Feminist science teachers claim that discrimination against girls in science is probably the rule rather than the exception.

> I feel that in some respects it is very difficult for men physics and chemistry teachers not to discriminate against the girls in their classes, even when they are aware of the possibility of this and have made a conscious decision not to do so.
>
> Samuel (1981, p.252)

3.1.3 Discussion

The quotes in the previous section certainly suggest that some science teachers do hold greater expectations for the boys in their classes than for the girls. These different expectations, on occasions, lead to differential treatment of boys and girls. "There is no shortage of anecdotal evidence from women who remember prejudiced science teachers, but there is a shortage of systematic investigation" wrote Alison Kelly in 1978b. Such is still the case.

Back in 1974, Mischel charged that "The scope and nature of sex bias have been subjected more to polemics than to objective research." A similar sentiment was expressed by Deaux and Taynor (1973), "Considerable verbiage but not much research may be found on the topic of equality between the sexes and, in particular, on the question of whether a woman is rated on a par with a man for a given level of accomplishment." Both of these quotes, which referred to psychology research in general, were written over ten years ago. During the intervening years, a number of studies have investigated aspects of sex bias in evaluation (see the work reviewed in section 2.5.1).

The problem of differential evaluation in educational settings was raised in 1968 by Garai and Scheinfeld in their conclusion to a lengthy article on intellectual and psychological sex differences. They wrote:

> ... where conventional school tests in science or mathematics are slanted toward the 'masculine' method of evaluating performance, the discrimination may be directed against the girls. Much research is still needed to provide more information about conscious and unconscious biasses in the evaluation of the performance of either sex by the same and the opposite sex ... (p.278)

Their call for more research into the sex bias of teachers was largely ignored. Certainly the marking practices of science teachers were not examined for possible sex bias. In 1982, whilst reviewing a paper on sex stereotyping in schools that had been written with science teachers in mind, Minden and Duelli-Klein concluded "We hope that this paper will stimulate increased attention to the sexism of educators, and inspire further strategies for exposing and doing away with such discrimination." It was the recognition and acknowledgement that such work was long overdue that inspired this research.

3.2 HYPOTHESES

3.2.1. Introduction

As was intimated in Chapter 1, large portions of this study are essentially exploratory. Little is known about science teachers' attitudes towards girls and boys, their expectations for each sex, and their reactions to girls and boys studying science. So a fundamental aim of this study is to elucidate and describe 'what is'. Since no material has hitherto been published on some of the topics chosen for investigation, it was often not possible to propose definitive hypotheses. Only on occasions was it possible to extrapolate from related research to produce meaningful hypotheses. In the absence of hypotheses, questions were used to direct and guide the research. A few of the topics investigated in this study have been researched in the past. The hypotheses associated with these topics tend to be more rigorously defined since they draw upon previous findings and existing theory.

All of the hypotheses and most of the questions guiding this research were informed by existing theory and/or the findings of previous studies reported in Chapters 1 and 2. The research as a whole is structured by a loose framework, described in Chapter 1. This model was also built upon existing facts and ideas. The specific presumptions that directly moulded the model, and indirectly shaped many of the hypotheses and questions, are outlined in Chapter 1.

The emphasis of the following hypotheses, and consequently of the whole study, is not on restructuring theory, but on clarifying and describing phenomena and relationships between them that have so far only been hazily glimpsed.

3.2.2 Sex typing of science

'Science teachers perceive science to be masculine'

Work with school children and students has shown that they view the physical science subjects as masculine subjects, and subjects that are more suitable for boys than girl. Physics is generally rated more masculine than chemistry (section 2.1.1). Biology contrasts with the physical sciences because it is viewed as neutral or even slightly feminine (section 1.3). If it is assumed that pupils' ideas about science mirror the views of adults around them, then two hypotheses can be proposed.

Hypothesis One Science teachers perceive physical science to have a masculine image.

Hypothesis Two Science teachers rank the three common science subjects in order of masculinity - physics, chemistry, biology - with physics being the most masculine subject.

The origin and meaning of the term 'masculine image' has not yet been adequately researched, so it was decided to investigate the following questions.

Question One How do science teachers account for the masculine image of science?

Question Two What subject characteristics are associated with a masculine image?

To determine the universality of science teachers' perceptions of science as a masculine subject, the following questions were asked.

Question Three How do the views of secondary science teachers about the masculinity of science compare with those of teachers of other subjects and teachers from other educational levels?

Question Four Do science teachers with different subject specialities vary in their views about the masculinity of science?

Question Five What personal and educational variables are associated with extreme views about the masculinity of science? (See also section 3.2.6)

If science teachers regard science as a masculine subject, then they probably expect people engaged in scientific study and research to be male. Furthermore, they probably expect those males to display very stereotyped masculine traits. These arguments gave rise to another hypothesis.

Hypothesis Three Science teachers mainly associate scientists with masculine qualities and rarely with feminine qualities.

3.2.3 Sex stereotyping

'Science teachers perceive differences between boys and girls which could affect science achievement'

(A) Research has shown that teachers hold sex stereotyped ideas about the behaviour of their pupils, the personality characteristics of their pupils, and the work related attributes of their pupils (see section 2.2.2). As regards work related attributes, boys are perceived to be more original, logical, analytical; whereas girls are more intuitive, industrious, conscientious, neat. If science teachers also stereotype the work related attributes of their pupils, then they are likely to associate the written work of boys and girls with different set of characteristics.

<u>Hypothesis Four</u> Science teachers recognise differences between the written work of girls and boys.

If the differences that science teachers perceive between the written work of boys and girls are sufficiently great and sufficiently obvious, then teachers may believe that they can tell whether a piece of written work was produced by a boy or a girl.

<u>Question Six</u> Do science teachers believe that they can generally distinguish between the written work of girls and boys?

Besides simply determining whether science teachers believe there are differences between the written work of boys and girls, the exact nature of those perceived differences was also investigated.

<u>Question Seven</u> What features do science teachers associate with the written work of girls and boys?

To determine whether science teachers' perceptions of girls' and boys' written work differ from those of teachers of other subjects, three further questions were asked.

<u>Question Eight</u> Do similar proportions of science teachers and other teachers recognise differences between the written work of girls and boys?

<u>Question Nine</u> Do similar proportions of science teachers and other teachers believe that they can distinguish between the written work of girls and boys?

<u>Question Ten</u> Do science teachers and other teachers associate similar features with the written work of girls and of boys?

(B) Since teachers' perceptions of pupils' personality traits, scholastic aptitudes, and interests are sex stereotyped, it is very likely they also believe that boys and girls prefer different characteristics in school subjects, and so are attracted to different types of subjects. Such reasoning led to the following hypothesis.

<u>Hypothesis Five</u> Science teachers believe that boys and girls have different preferences regarding subject characteristics.

To determine whether science teachers' beliefs about the subject characteristics preferred by boys and by girls are typical of the beliefs of teachers in general, the following question was asked.

Question Eleven How do secondary science teachers' perceptions of pupils' preferences for different characteristics compare with those of teachers of other subjects and teachers from other educational levels?

(C) In addition to sex stereotyping pupils' personality traits, scholastic aptitudes and interests, teachers also sex stereotype the roles that they expect their pupils to occupy in adult life (see section 2.2.2). The sex stereotyping of both present attributes and future roles is probably inevitable and also mutually reinforcing. The work reviewed in section 2.2.2 indicates that, inspite of small changes recently, teachers still hold fairly traditional attitudes towards the sex roles of adults. It was anticipated that the attitudes of science teachers are not significantly different from those of other teachers.

Hypothesis Six Science teachers believe in traditional sex roles.

The general public's attitudes towards sex roles have been studied extensively in recent years. A number of such studies have identified variables that are related to the attitudes held by a person. Commonly mentioned variables include a person's sex (Jean & Reynolds, 1980; Spence et al., 1973), age (Jean & Reynolds, 1980; Spence et al., 1973), marital status (Jean & Reynolds, 1980), social class (Parry, 1983), education (Fransella & Frost, 1977; Hall & Frederickson, 1979), culture (Fransella & Frost, 1977), and mother's employment (Stewart & Winter, 1974).

The evidence in support of the relationship between two of the variables, sex and age, and sex role attitudes has been very consistent, and has often arisen from studies with similar experimental designs to that of the present study. Therefore these two variables were chosen to guide the formulation and content of two further hypotheses.

Hypothesis Seven Male science teachers hold more traditional attitudes than do female science teachers.

Hypothesis Eight Older science teachers hold more traditional attitudes towards sex roles than do younger science teachers.

For another two variables, social class and maternal employment, the experimental design of previous studies differed from that of the present study, and so previous findings were only used to suggest questions for investigation.

The work of Parry (1983) showed that a group of working-class people held more traditional views of sex roles than did a group of middle-class people. Teachers are all classified as middle-class, so it was not possible to investigate the effect of their current social class upon their sex role stereotypes. Instead, it was decided to investigate the effect of their parents' social class, i.e. the social class milieu surrounding them during their formative years.

Question Twelve Do science teachers from a working-class background hold more traditional attitudes towards sex roles than teachers from a middle-class background?

Research with children has shown very clearly that the children of working mothers perceive sex roles in a less stereotyped and less traditional way than do the children of mothers who do not work (Marantz & Mansfield, 1977; Robb & Raven, 1981). Although evidence for a relationship between the sex role stereotypes of adults and their mothers' employment during their childhood is less abundant, nevertheless such a relationship must be inferred.

Question Thirteen Do science teachers, whose mothers were full-time housewives during their childhood, hold more traditional attitudes towards sex roles than teachers whose mothers were engaged in paid employment?

Two of the other variables, education and culture, were judged to be inappropriate to the present study, since the teachers under investigation were expected to display comparatively small differences on these two variables.

(D) Since teachers anticipate different adult roles for their male and female pupils, it is likely that they also regard some subjects to be more important for boys and others to be more important for girls. Subjects such as technology and home economics are particularly likely to be differentially valued for boys and girls. Such considerations suggested the following hypothesis.

<u>Hypothesis Nine</u> Science teachers believe that not all school subjects are of equal importance for boys and for girls.

Furthermore, if science teachers view science as a masculine subject, and believe that most science related jobs are performed by men, then they may consider science to be a more important subject for boys than for girls.

<u>Question Fourteen</u> Do science teachers consider the science subjects to be of greater value and relevance to boys than to girls?

3.2.4 <u>Attribution patterns</u>

'Science teachers attribute boys' and girls' success/failure in science to different causes'

The literature indicates that girls are more likely than boys to attribute their failure in sex typed subjects to lack of ability. Adults also perceive that the successes and failures of males and females are due to different causes, especially when the task is more closely associated with one sex than the other. The success of a male or the failure of a female on a masculine task is attributed to stable internal factors since the outcome is anticipated. In contrast, the failure of a male or the success of a female on such a task is attributed to unstable external causes since the outcome is unexpected.

No work has been located which specifically investigates the attributions that teachers make for their male and female pupils. However, Burger et al. (1982) report that the success of pupils who are expected to do well and the failure of pupils expected to do poorly are more often attributed to stable internal factors than are the opposite outcomes. Unexpected outcomes are more likely to be attributed to unstable factors.

If it is assumed that science teachers regard the physical science subjects to be masculine subjects, then attribution theory offers four hypotheses.

Hypothesis Ten Science teachers are likely to attribute boys' success at science to stable internal factors, e.g. ability.

Hypothesis Eleven Science teachers are likely to attribute girls' success at science to unstable factors, e.g. effort or luck.

Hypothesis Twelve Science teachers are likely to attribute boys' failure at science to unstable factors, e.g. lack of effort or bad luck.

Hypothesis Thirteen Science teachers are likely to attribute girls' failure at science to stable internal factors, e.g. lack of ability.

If teachers attribute boys' and girls' successes and failures in science to different causes, then it is likely that they also believe that boys and girls are differently motivated when they choose either to continue studying science or to drop science at 14+. No previous research into this topic has been located, and so no hypothesis could be made. Instead the work was guided by the following question.

Question Fifteen Do science teachers believe that boys and girls are differently motivated when they choose either to continue studying science or to drop science at 14+?

3.2.5 Teacher expectation and judgement

'Pupil's sex influences teacher's expectations and judgements' Science teachers who consciously or unconsciously believe that the physical sciences are masculine subjects, may also perceive girls as less suited for science studies than boys. These views could result in teachers holding higher expectations for boys' prospects and success in science. The expectations of teachers might also be reflected in the subjective evaluation of pupils' classroom performance and work submitted for assessment. Work of a high standard which has been produced by a girl might be assessed harshly because it is inconsistent with a teacher's preconception that science is a subject for boys, whereas comparable work produced by a boy might be overrated as it conforms to the teacher's expectations.

The research discussed in sections 2.4 and 2.5.2 indicates that teachers do sometimes hold different expectations for certain groups of pupils and do sometimes make biased judgements when marking. The few studies that have been conducted with science teachers (see section 3.1.1) reveal that they are no different from other teachers. Such observations suggested two hypotheses.

<u>Hypothesis Fourteen</u> For identical written work, science teachers award higher marks to boys than to girls.

<u>Hypothesis Fifteen</u> Based on the evidence of written work, science teachers form higher expectations for boys than for girls, as signified by their judgement of pupils' potential for science.

If science teachers hold higher expectations for boys and judge the work of boys more favourably than that of girls, then they probably also believe that boys are better suited to science studies, i.e. their cognitive and affective attributes are superior to those of girls. This supposition can be expressed as another two hypotheses.

<u>Hypothesis Sixteen</u> Based on the evidence of written work, science teachers are more likely to judge that a boy, than a comparable girl, possesses cognitive ability that is appropriate for the study of science.

<u>Hypothesis Seventeen</u> Based on the evidence of written work, science teachers are more likely to judge that a boy's attitude towards science and his interest in science are superior to those of a comparable girl.

Whilst comparing the marks awarded to boys and to girls by science teachers, there was an opportunity to compare the marks awarded by male and female teachers as well. Research evidence is inconclusive as to which sex generally awards the higher marks (see section 2.5.2.1.1).

However, the previous experimental study of science teachers' marking patterns by Spear (1982) indicated that female science teachers award higher marks than male science teachers. This particular finding was taken into account when wording a hypothesis.

> Hypothesis Eighteen When marking identical samples of written work, female teachers give higher marks than do male teachers.

Besides simply investigating the effect of pupil sex and teacher sex individually upon the marks awarded, there was also the opportunity to investigate the existence of any interactions between the two variables.

> Question Sixteen Does teacher sex interact with pupil sex to further complicate the marks awarded to pupils?

When marking, it is unlikely that all science teachers differentiate between the work of boys and girls to a similar extent. Information about teacher characteristics that are associated with highly sex biased marking patterns and with unbiased marking patterns would be very helpful, and so the following question was asked.

> Question Seventeen What personal and educational variables distinguish teachers who award very similar marks to boys and to girls from those who award very dissimilar marks?
> (See also section 3.2.6)

Finally, the question of whether pupil sex always elicits sex differentiated expectations and sex biased judgements from teachers requires consideration. The phenomenon of sex differentiated expectations was chosen for detailed study. Some studies have shown that when information about an imaginary person is scant, subjects make use of stereotyped beliefs to arrive at perceptions and judgements of the person or the person's behaviour. The availability of more information decreases the subjects' reliance upon stereotyped beliefs (Delia, 1972; Locksley et al., 1980; Rosen & Jerdee, 1974b). Presumably the effect of the stereotyped beliefs operates via biased expectations, which are then modified by the amount of information available. This supposition raised the following question.

Question Eighteen Do science teachers form more sex differentiated expectations for pupils when less information is available?

3.2.6 Relationships between variables

The work of other researchers has already provided indications of interconnections between most of the five broad topics under investigation, although not specifically for science teachers. The links between sex stereotypes and attribution patterns (Garland & Price, 1977; Post, 1981), attribution patterns and expectations (Frieze, 1980), and teacher expectations and teachers judgements (Cahen, 1966; Chase, 1979) have all been studied in different contexts and with different samples. The link between sex stereotyped perceptions/attitudes and sex differentiated expectations has not been invistigated, but the longer link between sex stereotyped perceptions/attitudes and sex differentiated assessments has (Holahan & Stephan, 1981; Sharp & Post, 1980). The only links that have not been studied at all are those of sex typing of science - attribution patterns, and sex typing of science - teacher expectations.
This study sought to provide some evidence regarding the following questions.

Question Nineteen Are those science teachers who regard the physical science subjects as masculine subjects, more likely to use different reasons to explain the successes and failures of boys than of girls?

Question Twenty Are those science teachers who regard the physical science subjects as masculine subjects, more likely to hold higher expectations for boys studying science than for girls?

Question Twenty One Are those science teachers who use different attributions to explain the successes/failures of boys and of girls, more likely to hold higher expectations for boys studying science than for girls?

Question Twenty Two Are those science teachers with sex stereotyped perceptions and traditional sex role attitudes, more likely to use different reasons to explain the successes and failures of boys than of girls?

Question Twenty Three Are those science teachers with sex stereotyped perceptions and traditional sex role attitudes, more likely to hold higher expectations for boys studying science than for girls?

Question Twenty Four Are those science teachers who hold higher expectations for boys studying science, more likely to assess the work and attributes of boys more favourably than those of girls?

Besides investigating relationships between the main topics themselves, this study, as already indicated, also sought to determine the relationships between a number of independent variables and three of the topics - sex typing of science, sex stereotyping and teacher judgement. These topics were chosen because they occupy boundary positions in that portion of the model under investigation (see Figure 1.1). A number of the independent variables chosen for investigation have already been mentioned in section 3.2.3 above. Additional variables were mostly chosen because they seemed to be superficially related to variables studied by other researchers (see 3.2.3). Thus teaching experience was chosen because it generally correlates with age, teacher status was chosen because it is reminiscent of socio-economic status, and various aspects of a teacher's past and present educational experiences were chosen because they relate to amount of education. A teacher's involvement in teaching compulsory science courses to 16+ was also investigated. It was anticipated that teachers who taught or had taught in schools where girls were not given the option of discontinuing their science studies, may view science as less masculine, and thus perceive and judge girls differently to teachers who taught in schools where many girls drop most of their science subjects at 14+.

The variables listed for investigation were, in the first instance, chosen for their possible relationships with teachers' sex role stereotypes. However, in the absence of empirical or theoretical guidelines, the same list probably represents as good a selection of variables for investigations in association with the sex typing of science and teachers' sex differentiated judgements as any other selection.

CHAPTER 4

RESEARCH DESIGN

4.0 Contents

CHAPTER 4

TABLES

Table 4.1 Stages of the research

Stage	Chapter(s)
Assessing resources and constraints	(Appendix 4.3)
Developing problems and hypotheses	1,3
Specifying population and sample	1,4
Deciding on methods of data collection	4
Preparing measuring instruments	6
Piloting and revision of measuring instruments	6
Fieldwork	4
Data processing and analysis	4,7,8
Reporting of findings	7,8

4.1 INTRODUCTION

"Research design is the plan, structure, and strategy of investigation conceived so as to obtain answers to research questions and to control variance" (Kerlinger, 1973, p.300).

The plan of a research project is the overall scheme and it should include all the stages of the research process. Table 4.1 shows the stages of this study and the chapters in which they are described. A research investigation is usually structured by a guiding model. The structural model outlining the important variables and their interrelationships in this study has already been presented in Chapter 1. Strategy refers to the details of data collection and analysis, including the solving of associated problems. This aspect of the present study is reported in sections 4.5 and 4.6. A detailed consideration of sources of variance and their control appears in Appendix 4.1.

4.2 CHOICE OF RESEARCH METHODS

A variety of methods are available for use in educational research. The three major styles of research are ethnographic, survey and experimental (Open University, 1979). Ethnography involves the detailed study of small groups of people within a complex society. The emphasis is usually on forms of social interaction and the meanings which lie behind them. The survey method involves asking a sample of respondents a number of fixed questions under comparable conditions. The orientation of a survey may be either descriptive or explanatory. A descriptive survey is concerned with discovering the relative incidence and distribution of various attributes of a population. An explanatory survey aims to elucidate interrelationships between variables. Surveys can also be classified according to the method by which information is obtained, e.g. personal interview, postal questionnaire, telephone survey. An experiment is a research study in which all the possibly relevant independent variables are controlled and/or manipulated by the experimenter. This allows cause and effect relationships to be clarified and hypotheses to be tested.

The basic aims of the present study were concerned with (a) establishing whether secondary science teachers sex type the science subjects, (b) establishing the existence of and the extent of sex stereotyped views among science teachers, (c) demonstrating sex biased expectations and judgements in a marking exercise, and (d) investigating the personal and educational correlates associated with any sex bias and sex stereotyped views displayed by secondary science teachers.

It was considered that the different research objectives could best be investigated by using a variety of research methods. For the marking exercise, a field experiment similar to the one employed in a previous study (Spear, 1984) seemed to be most appropriate. Attitudes are most commonly measured by attitude scales (Nunnally, 1978). Therefore, questionnaires composed of self-report measures were deemed to be the most suitable method for obtaining data relating to the teachers' attitudes, views, beliefs and opinions. To guide the content and construction of these self-report inventories, knowledge of secondary school science teachers' expressed ideas and views regarding the topics under investigation was required. In-depth exploratory interviews could most effectively supply this information.

A full discussion of the advantages and disadvantages that led to the selection of the three chosen research methods (field experiment, exploratory interviews and questionnaires) is presented in Appendix 4.2. Postal questionnaires were chosen as the primary research method to collect the bulk of the data for the investigation. This decision was guided by a number of factors, which referred to resources available, external constraints imposed, and the need for large samples. The various factors are itemized in Appendix 4.3.

4.3 DESCRIPTION OF QUESTIONNAIRES

4.3.1 Overview

A number of questionnaires and scales were produced and used. This section describes the content and format of those questionnaires which, after rigorous development and teating, were used to gather the main body of data reported in this thesis. Individual scales which, although providing valuable data, did not appear in the final form of a questionnaire designed to be completed by teachers are not considered in this section. However, every scale used in the study, regardless of its longevity, is listed in Table 4.2. The table also shows where descriptions of each scale's development and/or use can be located in the text, where results from the scales are reported, and where those results are discussed.

Most scales and questionnaires used in the study were associated with the development and piloting of the three principal questionnaires: Bases of Individual Assessment (BIAS), Science Teachers on Science Subjects (STOSS), and Characteristics of School Subjects (COSS). These questionnaires contained a variety of scales designed to measure different aspects of the central topics under investigation. Table 4.3 summarises the scales contained within the final form of each of the three main questionnaires, and the topics that the scales refer to. The table also indicates where copies of the questionnaires can be found in the appendices, and in which section of the questionnaire each scale appears. Further descriptions of the structure and format of the three questionnaires follow in sections 4.3.2 to 4.3.4, whilst details of the development of the individual scales appear in Chapter 6.

4.3.2 Bases of Individual Assessment (BIAS)

Section A of this questionnaire was designed to collect factual information about the respondent. Questions covered such areas as

Table 4.2 Scales: Their Reportage

Scale	Section where reported			
	Description	Development	Results	Discussion
Adjective pairs	6.6.1.5		6.6.1.5	
Cards	6.9.2 App. 6.14	6.9.2.1 6.9.2.2	App. 6.15 App. 6.16	App. 6.15
Characteristics of science	4.3.3 App. 4.5	6.6.3	7.1.3 8.1.2.2	9.1.1
Females' social roles	4.3.3 App. 4.5	6.7.3	7.2.3.2	9.2.3
Gender of word pairs	6.6.2.1	6.6.2.1.1	App. 6.5	
Importance of subjects	4.3.4 App. 4.7	6.7.4	7.2.4 8.1.3.2	9.2.4
Marking exercise	4.3.2 App. 4.4	6.9.1	7.4 8.2/8.3	9.4
Masculinity index	4.3.3 App. 4.5	6.6.2	7.1.2 8.1.2.1	9.1.1
Name preferences	6.9.1.2	6.9.1.2	App. 6.12	
Opinions	6.6.4 App. 6.9	6.6.4.1	7.1.4	9.1.4
Personal details	4.3.2 App. 4.4	6.10.1	App. 4.8	

Scale	Description	Development	Results	Discussion
Preference for subject characteristics	4.3.3 App. 4.5/4.6	6.7.2	7.2.2 8.1.3.1	9.2.2
Reasons for choosing/dropping	4.3.3 App. 4.5	6.8.2	7.3.3 7.3.4	9.3.2
Reasons for success/failure	4.3.3 App. 4.5	6.8.1	7.3.1 7.3.2	9.3.1
School details	6.10.2 App. 6.17	6.10.2	App. 4.10	
School subject characteristics	4.3.4 App. 4.6	6.6.1	7.1.1	9.1.1 9.1.2
Scientist stereotypes	4.3.3 App. 4.5	6.6.5	7.1.5 8.1.2.3	9.1.5
Variables used in marking	6.9.1.4		6.9.1.4	
Written work of girls and boys	6.7.1 App. 6.10	6.7.1	7.2.1	9.2.1

Table 4.3 Structure of the three main questionnaires

Questionnaire	Appendix	Section	Scale	Topic
BIAS	4.4	A	Personal details	
		B	Marking exercise	Teacher expectation Teacher judgement
STOSS	4.5	A	Masculinity index	Sex typing of science
		B	Characteristics of science	Sex typing of science
		C	Preference for subject characteristics	Sex stereotyping
		D	Reasons for choosing/ dropping	Attribution
		E	Reasons for success/ failure	Attribution
		F	Scientist stereotypes	Sex typing of science
		G	Females' social roles	Sex stereotyping
COSS	4.6		School subject characteristics	Sex typing of science
			Preference for subject characteristics	Sex stereotyping

personal details (e.g. sex, age, socio-economic background), educational background, qualifications, past teaching experience, and current teaching conditions.

Section B comprised the marking exercises. The design of the experiment was based upon a 2x2 factorial design, with repeated measurements. The two factors were pupil sex and teacher sex, whilst measurements were repeated over three different levels of pupil ability. A complete factorial design (i.e. 2(sex of pupil) x 2(sex of teacher) x 3(ability of pupil)) was rejected since it would have entailed 12 cells and a correspondingly small number of subjects in each. An experimental design which included repeated measurements was more economical in terms of time and effort, and especially attractive when the securing of a large sample was problematical. Having decided to ask each teacher to mark the work of three pupils of differing ability, it was thought inadvisable to present work from three pupils of the same sex, as such an arrangement might have caused the teachers to question the motives and objectives of the investigation. Hence, the teachers were shown work from both sexes. The arrangement of pupil sex and pupil ability across the samples of work was systematically varied so as to eliminate any effects which might have arisen from the order of presentation.

Although science teachers were asked to evaluate the work of three pupils, they actually saw six pieces of work - two from each pupil. The samples were handwritten and all of them were easily legible. The teachers were informed that the work samples had been written by 12-year-old pupils of average ability, following a combined science course in a comprehensive school.

The first sample from each supil was a write-up of a practical experiment on distillation, and the second was a homework essay entitled, "What I think about science and scientists". The three write-ups covered a range of different experimental methods, different styles of presenting work, and different levels of attainment. One sample was conceptually accurate, another contained correct ideas but some of the experimental details were faulty, whereas the third sample was scientifically incorrect. All three homework essays contained statements indicative of both positive and negative attitudes about science and scientists.

The original work samples had been produced by both boys and girls. However, each pair of samples was presented to half of the teachers as being the work of a girl and to the remaining teachers as being the work of a boy. This arrangement was plausible since the samples would not be readily associated with either sex (according to the results reported in section 7.2.1). Pupil sex was denoted by attaching a fictitious name to each pair of work samples. The names were viewed equally favourably by teachers (see section 6.9.1.2).

Booklets were formed by stapling together the samples, an introduction and instructions. Although the sequence of sample pairs was fixed, rotation of the sample pairs ensured that each sample pair was seen in the first position by approximately one third of the respondents. Additionally, in half the booklets the first sample pair was written by a 'boy', whereas in the other half it was attributed to a 'girl'. Each booklet contained two sample pairs written by one sex, and one sample pair written by the opposite sex. All booklets sent to a single school conveyed identical information regarding the sex of the authors.

The introduction described the marking exercise as an investigation to compare the standards set by teachers, and also the importance attached to a variety of work and pupil characteristics. To validate these ostensible objectives, when the teachers were asked evaluative questions at the end of each sample pair, they were not offered any criteria to guide their responses.

The instructions requested that the teachers should not discuss or consult with each other. A covering letter further stressed the same point.

Fifteen evaluative questions followed each pair of work samples.

Eleven of the questions referred to the experimental write-ups and four to the homework essay. The first three questions, which were concerned with the overall standard of the write-ups, requested respondents to assess the work on a simple 3-point scale, indicate what mark out of 10 the work merited, and also what mark they would actually award it. The next seven questions examined a number of work variables which teachers claim they take into account when assessing written work (see section 6.9.1.4). The last five questions referred to the pupil's aptitude for science, attitude towards science, interest in science, and suitability for O level and CSE physical science courses as implied by the content of the work samples. The teachers indicated their responses to these last 12 questions on 5-point scales.

4.3.3 Science Teachers on Science Subjects (STOSS)

The purpose of the questionnaire was to determine whether science teachers associate school science and scientists with masculinity, and whether they perceive differences between boys and girls which could affect science achievement. Science teachers' attitudes towards women's role in society were also explored. In an introduction, the questionnaire was presented to respondents as a survey designed to record the views of science teachers on (a) the subjects they teach, and (b) their pupils. Whilst being broadly accurate, this introduction did not divert teachers' attention from the obvious contextual link between a pupil's sex and his or her motivation and preference for science. Respondents were therefore also told that the purpose of the investigation was to assess the topical question of teachers' attitudes towards compulsory science courses for all pupils up to 16.

The first page of the questionnaire requested brief particulars about the respondent, including their sex and their principal teaching subject. The latter piece of information was important, since in some sections of the questionnaire respondents were asked to reply with respect to their main teaching subject. Respondents were also asked not to confer with each other until the questionnaire had been completed. One reason for this request was to discourage respondents from discovering that the questionnaire appeared in two different formats. In one format some of the sections asked about girls, whereas in the other format the same sections asked about boys. Half of the subjects were given the 'girl' format and half were given the 'boy' format.

The questionnaire consisted of seven sections. Some of the sections were further sub-divided. Specific instructions appeared above each section.

The first section, section A, contained sets of scales which provided masculinity indices for the three main science subjects. Each index was determined from an identical series of four semantic differential scales. The scales explored a teacher's affective response to the science subjects as they are taught up to CSE/O level standard. Since all the scales carried gender connotations (see section 6.6.2.1.1), they specifically measured a teacher's feelings about the masculinity of science.

Section B investigated the characteristics of school science subjects as seen by science teachers. Respondents were asked to rate subjects on a number of adjectives which are commonly applied to science. By including 'masculine' in the list, it was possible to establish which subject characteristics are associated with masculinity. To cut down the amount of time spent on this section, teachers were only asked to rate two science subjects - physics and biology.

The third section asked teachers what characteristics they believe make a school subject attractive to 14-year-old pupils, the age at which most pupils are faced with choosing a number of subject options. The question was structured around a series of nine semantic differential scales, the bipolar adjective pairs having been chosen to highlight a number of the differences that teachers perceive between arts and science subjects (see section 6.6.1). To expose any sex stereotyping of pupils' preferences for subject characteristics, teachers had first to complete the semantic differential scales with respect to the preferences of a typical 14-year-old girl, and then repeat the exercise for a typical 14-year-old boy.

Section D investigated further teachers' perceptions of pupils' motivation regarding their rejection or acceptance of the science subjects when subject options are chosen. Teachers were asked to rate how frequently they believe each of ten factors influence pupils' decisions to choose or drop science subjects. The ten factors covered the reasons most commonly given by science teachers to explain why pupils continue or drop science subjects (see section 5.3.2). The factors, appropriately worded, were set out twice. The first block related to reasons for pupils choosing science, and the second block to reasons for pupils dropping science. Half of the teachers were asked to complete the two blocks with reference to boys, and the other half were asked about girls. In addition, each teacher completed this section with reference to his or her main teaching subject. Thus the replies were split by pupil sex (boy, girl) and by subject (physics, chemistry, biology).

The fifth section, section E looked at teachers' attributions of pupils' academic achievements in science, a topic closely related to pupil motivation. Respondents were asked to consider the reasons for pupils' academic success and failure in the science subjects up to CSE/O level standard. Two corresponding lists of twelve factors (appropriately worded) were presented to the teachers and they had to rate the importance of each reason to pupils' success and then to pupils' failure. The factors represented a selection of the reasons that science teachers commonly give to explain the success and failure of their pupils (see section 5.3.4). Those teachers asked to answer the previous section with reference to boys also answered this section with reference to boys. Likewise, the other teachers answered both sections D and E with reference to girls. Section E also asked teachers to consider their main teaching subject when replying. Therefore, like section D, the replies were split by pupil sex (boy, girl) and by subject (physics, chemistry, biology).

Section F explored whether science teachers hold stereotyped views of scientists. To simplify the section, only two scientists were investigated - a physicist and a biologist. Respondents were asked to judge the likihood of each scientist possessing a number of characteristics and to record their replies on a set of semantic differential rating scales. The adjectives referred to a number of personality traits that are stereotypically associated with scientists (see section 2.1.1). Since it is believed that these traits help to distinguish scientists from non-scientists and hence from most women, their study is directly related to the stereotyping of scientists as men and indirectly to the sex typing of science as a masculine domain.

The last section was a shortened form of the Attitudes toward Feminism (FEM) scale (Smith et al., 1975), and it assessed sex role stereotyping. Specifically, it investigated respondents' attitudes towards the role of women in society. The scale consisted of five items and respondents had to indicate the extent to which they agreed or disagreed with the sentiment expressed in each item.

4.3.4 Characteristics of School Subjects (COSS)

The main purpose of this questionnaire was to investigate teachers' views about the different school subjects as they are taught in secondary school up to CSE/O level standard. The results should help to establish the degree to which teachers at different levels and of different subjects, sex type a range of school subjects, particularly the science subjects.

The first page of the questionnaire, besides providing an introduction, explanation and directions, requested brief classification type details about the respondent, e.g. sex, age, principal teaching subject. The instructions were very detailed and described the approach to the task, the items to be considered and the use of the rating scales. Respondents were also asked not to confer with their colleagues until after the questionnaire was complete.

The main body of the questionnaire contained eight sets of nine semantic differential rating scales. Each set of scales referred to a different school subject but otherwise all the sets were identical. The adjective pairs comprising the rating scales, described a number of characteristics that teachers believe differ between arts and science subjects (see section 5.3.1). The adjective pair 'feminine-masculine' was included as a direct measure of the perceived gender of a subject, and to act as a referend for the other scales.

The combination and order of school subjects presented for rating was variable. There were two different combinations of school subjects, chosen with regard to the range and popularity of subjects currently taught in secondary schools. To ensure that all respondents indicated their views concerning the science subjects, both combinations of subjects included physics, chemistry, biology and mathematics. The position of the subjects within the questionnaire was varied to counter order effects. Each science subject could appear in any of the eight positions, so there were a total of eight different formats of the questionnaire.

The last part of the questionnaire sought to establish those characteristics that teachers believe make school subjects attractive to 14-year-old pupils, the age at which many pupils choose their subject options. Respondents used the same semantic differential scales as they had used to rate the separate school subjects. To find out if teachers' views regarding pupils' preferences for subject characteristics are sex stereotyped, teachers had first to complete the semantic differential scales with respect to the preferences of a typical 14-year-old girl, and then repeat the exercise for a typical 14-year-old boy. This scale also appeared in the STOSS questionnaire, section C.

Besides variations in the nature and arrangement of the school subjects presented for rating, the COSS questionnaire also varied slightly depending upon the respondents whom it was designed for. There were two main forms of the COSS questionnaire. One form was for secondary school teachers, and the other was for primary school and middle school teachers. The two forms differed slightly in the wording of the introduction and instructions. An additional variation was also developed for use with science teachers, see Appendix 4.7.

The COSS questionnaire for science teachers only asked them how they view the three science subjects and mathematics. No other subjects were presented. The scale investigating pupils' preferences for subject characteristics was included. An extra scale, which inquired into the relative importance of different subject areas, was also included. Respondents were asked to rate how important they think CSE/O level qualifications in a number of different subject areas will be to pupils in their future lives. To find out if teachers' views regarding the importance of subjects are linked with a pupil's sex, teachers had first to rate the importance of the listed subject areas with respect to girls' future lives, and then repeat the ratings for boys.

4.4 SAMPLES

4.4.1 General considerations

Random sampling controls extraneous independent variables, which may influence the dependent variables, without first having to identify them. This prevents the introduction of systematic bias into the selection procedure, and results in a representative sample. Representative samples are desirable since descriptive statistics relating to them can be viewed as estimates of the equivalent measures for the population.

This correspondence allows findings referring to the sample to be generalized to the whole population.

True random sampling was not practicable in the present investigation for a number of reasons. Firstly, it would have been virtually impossible to identify and list every science teacher in all the maintained secondary schools of England. Without a complete sampling frame simple random sampling (and also systematic sampling) is not possible. Single-stage or multi-stage cluster sampling were the most feasible alternatives to random sampling. The original intention had been to adopt a single-stage cluster sampling technique, and sample all science teachers in a random selection of secondary schools in England. Unfortunately not all of the schools approached were willing to cooperate in the investigation. Therefore the research had to rely upon opportunity sampling.

Opportunity sampling, although widely used in educational research, is the weakest form of sampling. Strictly, the techniques of statistical inference should not be used on the resulting data, but their use has become conventional (Engelhart, 1972). Such use obviously does not eliminate the biases which can result from working with non-random samples. Nevertheless, it is generally accepted that if a non-random sample does not differ appreciably on a number of known characteristics from the population, then it may be permissible to regard the sample as being representative of the population and therefore treat it as if it were a random sample (Ferguson, 1976). In this study, many of the sample characteristics investigated did not differ significantly from those of the whole population (see Appendix 4.8) and so a decision was taken to use inferential statistical methods. Further arguments supporting this decision are presented in section 4.6.1.

The optimal size of a sample depends upon the variability of the data being sampled, the type of sampling used and the degree of precision required. The sample must be large enough to be representative and most authorities advise 'Use as large samples as possible' (Kerlinger, 1973, p.127). With a large sample the sampling error is likely to be small, thus the statistics calculated from large samples are more accurate and stable than those calculated from small samples. In this study the aim was to achieve at least 20, and preferably 30 subjects in each sub-group for which a statistical analysis was likely to be performed. This meant that some samples were very much larger in order to ensure that rare sub-groups,e.g. female physics teachers, were sufficiently large. The preferred minimum sub-group size was set at 30, since this is commonly the lower limit set for large sample inferential statistical analysis, e.g. t test (Gregory et al., 1969). Attention should also be paid to the dangers of working with samples that are too large, as they allow trivial associations to show up as significant results (Hays, 1974). In the present investigation concern was always directed towards samples that were too small, rather than too large.

4.4.2 Educational establishments

4.4.2.1 Schools

A total of 214 schools throughout England were involved in the research. Local schools were used for the exploratory interviews, and some of the initial pilot work. Some of the preliminary studies were conducted in the schools of known contacts. The majority of the pilot work and all of the final survey work was carried out in schools specially contacted for that purpose, and which were previously unknown to the researcher. This sample of schools was chosen from 'The Education Authorities Directory and Annual'. There was no systematic basis for choice, only that the schools should vary on a number of characteristics, e.g. type, size, location. Although the majority of pupils at maintained secondary schools attend mixed schools, it was decided to include some single sex schools in the sample to increase the heterogeneity of certain teacher characteristics, and to allow comparisons to be made between teachers from mixed and single sex schools. Even though the schools were not selected on a rigorous random or quota basis, nevertheless there is no reason to believe that systematic bias occurred.

Brief details of the schools which participated in the research appear in table A4.9/1 (Appendix 4.9). The table includes information about the types of schools used, whether they were coeducational or single sex, and their geographical location. Additional information was received from some schools, especially those which returned BIAS and STOSS questionnaires. These extra details are recorded in Appendix 4.10, together with an appraisal of the representativeness of the schools included in the BIAS and STOSS sample.

4.4.2.2. University Departments of Education

Postgraduate Certificate of Education (PGCE) students from six University Departments of Education (UDEs) throughout England were involved in reliability studies. Most of the departments agreed to assist in the research because one member of staff was interested in the research problem. However, there is no reason to suppose that the participating departments are particularly unrepresentative of UDEs in general, or that the replies of the students would have been biased in any way.

4.4.3 Subjects

4.4.3.1 Teachers

Details of the number, sex and subject speciality of the teachers composing each sample appear in Appendix 4.11. Most of the samples consisted of at least 35 respondents, and both sexes were always represented.

Within each school or science department, it was the intention that all eligible teachers should complete a questionnaire. Unfortunately this was not always achieved. Tables 4.4 and 4.5 indicate the number of

Table 4.4 Number of BIAS returns per school

No. of returns	1	2	3	4	5	6	7	8	9	12	13	16
No. of schools	26	8	7	18	9	5	7	2	3	1	1	1

Table 4.5 Number of STOSS returns per school

No. of returns	1	2	3	4	5	6	7	8	9
No. of schools	32	8	15	12	6	2	3	1	1

BIAS and STOSS questionnaires returned from individual science departments. Although the return rate from some schools was low, there is no reason to suppose that a teacher's decision to complete a questionnaire or not was related to any of the variables under investigation. Thus it is argued that even though some of the samples of teachers were partly self-selected, the samples were still reasonably representative.

The question of the representativeness of the respondents is considered further in Appendix 4.8 with particular reference to the BIAS and STOSS sample. The BIAS and STOSS sample was selected for subjection to detailed scrutiny, since the results from this sample constitute the major part of this thesis, and are central to the research problems and hypotheses as set out in Chapter 3.

4.4.3.2 PGCE students

Details of the number, sex and subject speciality of the PGCE students who assisted with the reliability studies appear in Appendix 4.11. Because the samples are small (in spite of considerable efforts to secure the cooperation of more subjects) and largely self-selected, it is unlikely that they are representative of the total population of PGCE students. Even so, there is no reason to believe that the attitudes and responses elicited by the questionnaires would be biased in any way.

Besides querying the representativeness of the PGCE students, the nature and extent of their teaching experience could also be questioned. But since they were in their final term and had all spent several months in schools on teaching practice, it is argued that they were adequately qualified to answer the questionnaires. It should also be borne in mind that they could very well have been picked up in a school sample a few months later.

4.5 DATA COLLECTION

Two broad methods of collecting data were employed. The interviews and some of the pilot work involved personal visits to meet and speak with the respondents. In contrast, nearly all of the questionnaires (most of the preliminary and pilot questionnaires, and all of the final form questionnaires) were sent by post, and the respondents had no personal or direct contact with the researcher.

A detailed list of the ways in which questionnaires were administered to all the samples used in the research appears in Appendix 4.12. Additional information concerning the content of the questionnaires and the composition of the samples is also included in the appendix.

Questionnaires were administered to most samples by an agent. The advantage of using agents is that they generally improve the response rate by reminding forgetful subjects and encouraging unwilling subjects. Also they can be enlisted to return the questionnaires in a single batch. In schools, the agent was usually the head teacher or head of science, although assistant teachers helped with some of the pilot work. In UDEs, questionnaires were administered to PGCE students by a lecturer.

Contact with schools was generally initiated by sending a letter to either the head teacher or head of science outlining the purpose of the investigation and requesting the cooperation and assistance of their school or science department in the investigation. The letter also contained an assurance of confidentiality, and a promise to supply a summary of the findings. Questionnaires were subsequently sent to those schools whose head teacher or head of science expressed interest in the investigation and willingness to volunteer their staff as subjects. Some of the BIAS questionnaires were distributed in a slightly different manner. Questionnaires were sent to a few heads of science directly without first checking that they were willing to assist. A few other heads of science were contacted via a couple of LEA science advisers, who were particularly interested in the investigation and eager to help with the distribution of questionnaires.

STOSS questionnaires were only sent to sampling units or individuals who had already completed and returned BIAS questionnaires. It was necessary to separate both the content and the distribution of the BIAS and STOSS questionnaires, to prevent replies to the BIAS questionnaire being contaminated by the notions contained within the STOSS questionnaire. It was feared that the topics presented in the STOSS questionnaire and their mode of presentation, might alert respondents to the true nature of the investigation, or at least stimulate respondents to speculate about the possibility of covert aims and interests on the part of the researcher. Either of these outcomes could have seriously biased responses to the BIAS questionnaire.

All questionnaires sent out were accompanied by a covering letter and a pre-paid reply label. Generally, the covering letter expressed gratitude for the assistance offered, and stressed that respondents should not discuss the questionnaires until after they had been completed. Some covering letters, e.g. the one sent with BIAS questionnaires, emphasized the need for a high response rate, and also asked the agent to complete a School Details questionnaire.

Reminder letters were sent to agents who had not returned questionnaires after a month. These letters were mildly worded and simply solicited cooperation. They stimulated most agents to return questionnaires. However, second follow-up letters had to be sent to a few agents. These letters were more pressing. The acknowledged that teachers may be reluctant to fill questionnaires, but stressed the need for questionnaires to be returned. Hardly any agents failed to respond to this second plea.

Although practically all schools and science departments returned questionnaires, the response rate from individual schools and science departments was very variable. From some places it was 100%, whereas from others it was somewhat lower. The number of BIAS and STOSS questionnaires received from each science department has already been recorded in Tables 4.4 and 4.5. The returns for the different formats of the BIAS, STOSS and COSS questionnaires were broadly equivalent (see Appendix 4.13 for further details).

4.6 DATA ANALYSIS

4.6.1 Parametric and nonparametric statistical tests

Data can be analysed using either parametric or nonparametric statistics. Parametric tests are more powerful. This means that, for a given level of significance, they require a smaller sample size to detect a true alternative to the null hypothesis (Hays, 1974). However, parametric tests involve a number of assumptions about the nature of the distributions of the variables in the populations from which the samples are drawn. It is assumed that (a) the population from which the samples are drawn have normal distributions, (b) variances are homogeneous from group to group, and (c) the data relate to interval or ratio scales, that is scales which are continuous and of equal interval throughout. By their very nature, such assumptions are frequently untestable, so the use of parametric techniques of data analysis involves risks.

Nonparametric tests make fewer assumptions about distributions; they do not require a normal distribution nor equal group variances. Besides being appropriate for use with data that, although from interval scales, have grossly nonnormal distributions, nonparametric methods are also appropriate for nominal and ordinal data. Advantages of nonparametric techniques include their wide applicability to all sorts of population distributions, and their ease of computation. A major disadvantage of nonparametric techniques is that they are not appropriate for complex analyses in which a large number of variables are to be manipulated.

Moreover, nonparametric methods are less powerful than parametric ones. They are somewhat less likely to reject a null hypothesis when it should be rejected.

On account of the advantages of parametric methods over non-parametric methods, many education and psychology researchers prefer to use parametric statistics to analyse data that are strictly only at the ordinal level. Most persuasive arguments that the powerful, parametric statistical tools (e.g. 'F', 't' and 'r') should be used when analysing social science data have been put forward by Bohrnstedt (1970) and Kerlinger (1973). Certainly, social scientists who have adopted this approach have demonstrated that the results obtained from assuming interval data are satisfactory and fruitful. Because it is probable that most psychological and educational scales approximate to equal interval scales, Kerlinger suggests that researchers stand to lose much by refusing to use the powerful parametric tools of measurement and analysis, and they are left with inadequate tools to solve their problems. Even though measurement errors may occur from assuming interval measurement, generally the result of such errors is the weakening of relationships among the variables (Bohrnstedt). That is, the result will appear to be more attenuated than they are in reality. Thus, the decision to assume interval measurement when it does not exist is unlikely to lead to spurious over-estimation of results. Consideration should also be paid to the fact that many of the common statistical techniques can withstand some departure from their fundamental assumptions about the measurement scale of the data (Nunnally, 1978). For example, Baker et al. (1966) have shown that the t test can be used with data that does not meet the criteria of an interval scale.

The two assumptions of normality and homogeneity of variance have also been examined empirically. Artificial populations have been set up, samples drawn from them, and t and F tests performed. Even when the assumptions of normality and homogeneity are violated, the probability statements resulting from the use of t and F are usually still highly accurate (Boneau, 1960). After considering evidence from a number of workers, Kerlinger (1973) concluded that the importance of normality and homogeneity has been overrated; and recommended "In most cases in education and psychology, it is probably safer - and usually more effective - to use parametric tests rather than nonparametric tests" (p.315). Anderson (1961) concurred and wrote that parametric procedures are the standard tools of psychological statistics.

Having regard for the advantages of parametric methods, it was decided to follow the recommendations outlined above and to make full use of parametric tests in the present study. In view of the robustness of parametric tests (meaning that departures from the assumptions of interval data, normality and homogeneity have little effect upon descriptive statistics or the probabilities obtained from significance tests), the data were not tested for departures from normality, homogeneity of variance and equality of intervals. It was assumed that the data met the three criteria.

Other workers using similar experimental designs, e.g. the Goldberg design, and similar scale formats, e.g. the semantic differential, to those used in this study have analysed their data with parametric techniques. Likewise, workers investigating similar topics, e.g. attribution theory, also used parametric techniques. Thus the use of parametric statistical analysis in this study allows the results to be more easily compared with those from other workers.

4.6.2 Probability and significance

In reaching a decision about the significance of differences in any data, two types of error may arise. A type I error is that made when the null hypothesis is falsely rejected, and a type II error is that made by not rejecting the null hypothesis when it is false (Hays, 1974). If too strict a level of significance is adopted, the researcher may fail to reject the null hypothesis when in fact a fairly large difference between the two population means actually exists. Because of the nature of the present study and because it was largely exploratory, it was felt that the risk of erroneously accepting null hypotheses would be more of a disservice to girls in science than would be the risk of erroneously rejecting them. Therefore the significance level was set at 5%. This means that there is a relatively low risk of falsely accepting null hypotheses, but a higher risk of falsely rejecting them. If the findings of this study were intended to be used as a basis for building theories or advocating social action, then a higher level of significance would be required.

Significance levels specify the probability of a chance occurrence of findings for a sample. By setting the acceptable significance level at 5%, results with a probability equal to or less than 0.05 are deemed to be non-chance, meaningful and significant, whereas results with a probability of 0.06 are unquestionably rejected. However, probability levels of 0.05 and 0.06 are not greatly different (Rosenthal & Rubin, 1979). The cut-off point at 0.05 is purely arbitary. To convey more information regarding how probable a particular result is by chance, Lewis (1967) advocates recording the exact level at which each result is significant. In practice, probabilities are customarily recorded at four levels: $p < 0.001$ (very highly significant), $p < 0.01$ (highly significant), $p < 0.05$ (significant) and $p < 0.1$ (marginally significant) (Ferguson, 1976; Youngman, 1979). This convention is employed in the present study.

On occasions, the significance levels will be indicated by a system of stars as follows:

Stars	Level of significance
***	< 0.001
**	< 0.01
*	< 0.05
+	< 0.1

Little mention is made of $p = 0.1$ levels since these are below the acceptable significance level. However, there is some justification for not totally ignoring such results. If more results at the 10% level were obtained by another researcher investigating another group of similar teachers, the probabilities from the independent tests could be combined (Fisher, 1950; Wallis, 1942) to produce an overall value that may well be significant at the 5% level.

When similar patterns, which are either non-significant or significant but small in size, are repeated on several independent variables, a consistent underlying relationship is indicated. Any failure to reach statistical significance could be due to the relationship being weak, poor measurement devices, or small sample size.

The approach frequently employed in this research of investigating a research hypothesis or question by using several statistical tests is basically problematical. Interpretation of the results may be difficult because type I error probabilities accumulate. A commonly adopted solution is to decrease the significance level for each test or to use a multiple-comparison procedure (Keppel, 1973; Kirk, 1968). However, both strategies result in lower power for each test, i.e. in larger probabilities of type II errors. Westermann and Hager (1983) have demonstrated that high power for each test is more important than a low significance level when the research hypothesis is an alternative hypothesis. Since alternative hypotheses, rather than null hypotheses, predominate in this research, high power for each test was crucial and so no attempt was made to use multiple-comparison procedures to compensate for cumulation of type I error probabilities.

4.6.3 Directional and nondirectional tests

The probability associated with a statistical test depends upon the hypothesis or predictions made before the data are collected. If neither theory nor common sense predicts a specific outcome, then a non-directional or two-tailed test is performed. Thus a two-tailed test would be appropriate if interest was focused upon the magnitude of a difference, not upon the direction of the difference.

If definite hypotheses or common sense predicts the direction of a statistical result before the data have been gathered, then a directional or one-tailed test should be used. The probability associated with statistical significance for a one-tailed test is half of that for a two-tailed test. Thus if there is a real difference in the predicted direction it will be more readily detected using a one-tailed test. The use of one-tailed tests when appropriate is generally recommended (Engelhart, 1972; Ferguson, 1976; McNemar, 1962), although caution should be exercised. It is particularly stressed that the prediction concerning the direction of the results, and the decision to use a one-tailed test should be made before the data are gathered (Lewis, 1967; McNemar, 1962).

The advice that it is generally more prudent to base conclusions on two-tailed tests (Lewis, 1967; Nunnally, 1975) has been followed with the result that nondirectional tests have been used extensively in this study. One-tailed tests have only been used where directional hypotheses were made prior to collecting the data. The one-tailed tests are clearly identified in the text. All of the other tests were two-tailed.

4.6.4 Educational significance

The significance levels given in this study only give information about how probable a particular result is by chance. Statistical significance does not necessarily correspond with educational or social significance. However, it is helpful first to establish evidence about the reality of a difference. But its magnitude is determined by referring to the sample means, not the level of statistical significance, i.e. its 'p' value (Jolly & Gale, 1976). Then its educational importance or significance is assessed in the light of related facts and experiences. Thus, the interpretation of results and their implications for science education are greatly influenced by the society in which we live.

4.7 DESIGN VALIDITY

Validity is defined as the degree to which a researcher has measured what s/he set out to measure (Smith, 1975). Two types of validity can be distinguished - validity of findings and validity of measurements. The former type of validity is discussed in this section, whilst the latter type is dealt with in section 6.4.

The validity of the researcher's interpretation of his/her findings can be threatened by a number of plausible alternative hypotheses if they have not been controlled in the research design. Campbell & Stanley (1963) have set out lists of possible sources of invalidity. They divide these extraneous variables into two groups, those concerned with internal validity and those concerned with external validity. Internal validity refers to whether what is interpreted as the 'cause(s)' actually produced the 'effects' in a piece of research. External validity is concerned with whether the results of the study can be generalized to other groups of people in other situations.

Campbell (1969) lists nine threats to internal validity. Because each questionnaire obtained data from each respondent on one occasion only, and because respondents were not assigned to the different questionnaire formats on a systematic basis, but on a more random basis, five of the nine variables do not apply to the present study or are exceedingly unlikely to have been operable. They are history, maturation, statistical regression, experimental mortality and selection-maturation interaction effects. A sixth variable, instrumentation, can also be discounted since both the questions asked and the scoring of the responses remained stable over time (see also Appendix 7.4). The three factors relevant to the present study are testing, instability and differential selection of respondents.

Any measurement that requires respondents to partake in unusual activities is likely to be reactive, i.e. to affect the respondent. In the marking exercise the teachers were requested to perform an activity which is very much a part of their working lives, and so testing effects should have been minimal on the BIAS questionnaire. The STOSS and COSS questionnaires were more susceptible to testing effects. However, it should be noted that testing effects most often become a serious threat to internal validity when respondents are subjected to a pre-test before the main test (Tuckman, 1978). Since this was not the case in the present study, the reactive effects of the STOSS and COSS questionnaires were probably random and small in magnitude.

The second factor, instability or chance fluctuations, is always present in any measurement. Statistical tests of significance are used to assess the plausibility of the data in the present study. If this leads to acceptance of the initial hypotheses, in spite of a risk that the decision is wrong, then the study will suffer from type II instability errors. However, if the risk of type II errors is decreased, then the risk of type I errors will increase. The researcher can only aim for the most appropriate balance between type I and II errors.

Differential selection of respondents, the third factor, was the one most difficult to control. The original intention had been to use cluster sampling to draw the sample, but in fact the participating schools were largely self selected, and so were individual teachers whithin each school. In spite of the questionable sampling technique, there is no evidence to indicate that the sample was not representative of science teachers in comprehensive schools. Their sex ratio, teaching qualifications and subject distribution were certainly typical of the whole population of secondary school science teachers (see Appendix 4.8). Furthermore, as regards the BIAS questionnaire, the true nature of the experiment was so well disguised, that even though there may have been differential selection of respondents, it was certainly not on the basis of their views regarding the suitability of science for boys and girls.

Campbell (1969) lists six possible threats to external validity. Because the measurements made in the present study relied upon a single approach to respondents, and most of the measurements have not yet been repeated, three of the six variables are not applicable. They are reactive effects of testing (pre-test and post-test sensitization), multiple treatment interference and irrelevant replicability of independent variables. The three factors that could threaten external validity are interaction effects of selection bias, irrelevant responsiveness of measures and reactive effects of data collection arrangements.

Samples differentially selected from the larger population may give responses unrepresentative of that population. In the present study, replies were received from a range of schools differing in size, type, location and socio-economic catchment area. Furthermore, the schools were scattered throughout England. Therefore the teachers in those schools adequately represent the target population.

Measures are complex and they always include irrelevant components that may produce spurious effects. Careful development of the scales used in the present study should have minimised these irrelevancies.

Reactive effects of data collection arrangements were seen to be the main threat to external validity. These effects are similar to testing effects, but differing emphasis is given to their effect upon validity. Testing effects threaten the interpretation of the findings, whereas reactive effects threaten the generalizability of the findings.

The knowledge that they are participating in a study, may lead some respondents to change their behaviour. They may desire their responses to reflect favourably upon them, or they might adapt their responses to match their perceptions of the researcher's expectations, even though the researcher may not have stated any objectives or intentions. In the present study, the aims of most of the investigations were disguised. To check that respondents had not surmised these objectives that were not stated, they were encouraged to write comments upon the aims and/or design of the investigation. Orne (1962) suggested that this technique indicates a respondent's perceptions of the experimental hypothesis. Many comments, often very lengthy ones, were received from respondents, but there were no indications that the respondents had ideas about the investigations that would invalidate the results. This was particularly true of the BIAS respondents. None of them appeared to have guessed the real aim of the investigation. Thus it is highly unlikely that external validity had been threatened by reactive effects resulting from the data collection arrangements.

Although internal validity could have been better, there is no reason to suppose that it seriously threatened the validity of the results. Alternative explanations or interpretations of the results can almost certainly be ruled out. External validity also appears to be satisfactory. This means that the findings of this study can confidently be generalized to other groups of science teachers, so long as the type of marking and the context of the marking is comparable to that of this study.

CHAPTER 5

EXPLORATORY INTERVIEWS

5.0 Contents

5.1 OUTLINE

The main purpose behind conducting exploratory interviews was to gain an initial insight into science teachers' views about the 'Girls and Science' problem. By probing their understanding of, and their explanations for the problem, it was anticipated that many salient variables would be identified. (This point is discussed further in Appendix 4.2. Knowledge about these primary variables, and the commonly perceived interactions between them, was needed to guide the development of measuring instruments that could subsequently be used to determine teachers' attitudinal and perceptual stances. In addition, the interviews provided an opportunity to confirm and clarify some commonly accepted ideas about school science and scientists.

A total of 26 secondary school science teachers were interviewed. However, three of the interviews were not recorded verbatim, and so the responses recorded in this chapter derive from 23 teachers (sample INT). Further details about the respondents and the schools in which they taught can be found in Appendices 4.9 and 4.11.

Most of the interviews were conducted individually, although on two occasions a couple of teachers were interviewed together. The interviews were loosely structured around a collection of 19 questions (reproduced in Appendix 5.1). These questions were merely intended to indicate the major topics to be broached in an interview. Not all of the questions, or even all of the topics, were necessarily covered in each interview. Furthermore, the order of introducing the topics was kept flexible, dependant upon the direction of the discussion. On occasions, topics raised by the respondent were discussed further and in greater depth.

Details regarding the recording of the interviews and the coding of the resulting data are presented in section 5.2. A broad summary of the findings is presented in section 5.3. More detailed results, including a quantitative analysis, can be found in Appendix 5.2. The information contained in both section 5.3 and Appendix 5.2 has been organised into the same sub-sections to aid interpretation and comparisons between the two complementary summaries of findings. The sub-sections correspond to the six broad topics that were discussed in most of the interviews. They were: (a) school subjects, (b) pupils' science choices, (c) the masculine image of science, (d) causes of success and failure at science, (e) differences between girls and boys, and (f) scientists. Discussion of the findings presented in section 5.3 is deferred until Chapter 9. It is then included with a discussion of the findings obtained from scales that ensued from the interview data.

5.2 RECORDING THE DATA

The whole of each interview from each sample INT teacher was tape recorded. This method of recording data is recommended by Engelhart (1972). It has the advantage of providing a complete record of what was said, thus eliminating bias due to conscious or unconscious selection by the interviewer of what to record. This aspect of tape recording was considered to be especially beneficial in the present study. The main disadvantage of tape recording becomes apparent with respondents who are very shy or inarticulate. However, such problems did not arise in the present study. All of the respondents were quite willing to have their replies taped and there were no indications that the presence of a tape recorder affected those replies.

After the interviews had been transcribed, the information was summarized using a method suggested by Henerson et al. (1978). For the first transcription, all of the relevant points mentioned within the six broad topics under investigation were recorded briefly. Also a tally mark was recorded for each point. As subsequent transcriptions were analysed, new points were recorded as described above. When a point that had been mentioned before was encountered, another tally mark was added to the original summary. Once all the transcriptions had been analysed and all the different points mentioned by the teachers had been recorded, then the raw data had to be classified. A manageable number of categories were defined and the raw data tallies were assigned to the different categories. The construction of the categories was guided by a set of recommendations proposed by Guilford and Fruchter (1973). They advocated that categories should be (a) clearly defined, (b) mutually exclusive and independent, and (c) exhaustive.

5.3 FINDINGS

5.3.1 School subjects

Replies to queries about possible ways in which school subjects and science subjects can be divided dichotomously, confirmed that the majority of teachers still (a) divide school subjects into science subjects and arts subjects, and (b) divide science subjects into physical science subjects and biological subjects. All of the respondents, except one, gave physics, chemistry and biology as examples of science subjects. Other examples mentioned by more than one quarter of the respondents were maths, general science, technology, geology and geography. The most frequently mentioned examples of arts subjects were history, geography, English, modern languages and art.

In reply to a question about differences between science subjects and arts subjects, the most commonly mentioned point was the observation that the science subjects are practical subjects.

> Science subjects are practical subjects; lots of time spent in practical work; more structured; rely more on the analytical perhaps, than the creative aspects of the mind. Although I wouldn't like to say that science isn't creative; I think it is in certain situations: In school, science tends to be more analytical.
>
> Male chemistry teacher (1)

A number of points, besides the practical component of science, are mentioned in the above quote. The analytical side of science was mentioned by several respondents.

> They require a slightly more analytical viewpoint. ... You have got to look at something and draw conclusions from it. In the arts subjects I tend to feel that the information is presented en masse

> and you just absorb it as it stands. Science involves elements of measurement and analysis which you don't find in other subjects.
> Male general science teacher (13)

Other respondents used the term 'logical' to describe the thought processes associated with science.

> They are much more logical in approach, whereas often with the humanities they don't appear to be teaching logical thought. Also in science you are tending to deal with facts rather than opinion.
> Female biology teacher (12)

The factual nature of science was mentioned by many teachers.

> It's more factual and structured.
> Female physics teacher (15)

> Essentially it is learning facts which have been established and which have become accepted, and learning skills and techniques which have developed. Whereas arts subjects are really open-ended and you are encouraging the pupils to present pieces of work which have not been established previously ... it's creative. I suppose creativity is the basic difference.
> Male physics teacher (9)

A few teachers referred to the science subjects as being highly structured subjects (see the quotes from teachers (1) and (15) above), whereas the arts subjects were viewed as being open-ended or unstructured (see the last quote above). Several teachers mentioned the lack of opportunity for creative work in the science subjects (see the quotes from teachers (1) and (9) above).

A few respondents referred to the objective-subjective continuum, placing the science subjects firmly at the objective end.

> I would have thought that they are often more objective rather than subjective. Most of science presumably is based on thesis and hypothesis and application of principles, whereas many of the so-called arts subjects, I'd imagine, are more drawing conclusions from feelings.
> Male chemistry teacher (14)

Differing emphasis on numbers and words between science and arts subjects was mentioned by four teachers.

> The use of maths and the vocabulary that we use. I think you need a more extended vocabulary for arts subjects.
> Male physics teacher (5)

Perhaps the mathematical content of science accounts for the last point.

> They (science subjects) are harder.
> Male general science teacher (13)

In summary, there was general agreement that the science subjects are more practical, factual, analytical, logical, objective and structured than the arts subjects. Specific mention was made that the arts subjects are more creative, opinionative and subjective than the science subjects.

5.3.2 Pupils' science choices

Three questions asked teachers directly for their views about the factors that influence pupils to choose or drop science. The questions were:

(i) Why do more girls than boys study biological subjects?

(ii) Why do more boys than girls study physical science subjects?

(iii) Why do most girls drop physical science when choosing their subject options at 13+ or thereabouts?

The discussions stimulated by these three questions contained the respondents thoughts about four obvious trends.

(i) Why girls choose biology.

(ii) Why boys reject biology.

(iii) Why boys choose physics.

(iv) Why girls reject physics.

An additional question asked teachers for their views as to how more girls could be encouraged to study the physical sciences. Since this concern goes beyond the topics under immediate investigation in the present study, the teachers' answers are reported in Appendix 5.3.

5.3.2.1 Girls choose biology

The single factor mentioned by most teachers to explain why girls choose biology was that of interest.

> It's just a matter of general interest. Girls are particularly keen on human biology. I suppose they relate this to their own future as mothers.
>
> Male general science teacher (11)

The category into which interest was classified, the category of affective factors, was also the most frequently mentioned category. The other main factor that appeared in the category was the feminine image of biology.

This factor was mentioned by several respondents to explain girls' choice of biology.

> It's what is loosely called a girls' subject. Don't ask me why it is called a girls' subject ... I don't honestly know. I have never thought about it too much.
>
> Male physics teacher (8)

Many respondents believed that social factors and social pressures were largely responsible for influencing girls to choose biology.

> It's tradition. We have come, I don't mean we as teachers, I mean as a society, we have come to expect that girls will go on the life sciences and boys will go for the physical sciences. I think that it's because we expect it, because they have come to assume that we expect it, that they do this.
>
> Male general science teacher (13)

The particular relevance of biology to a girl's future role in establishing a family was also mentioned by a number of teachers.

> I think they feel biology has more immediate relevance to themselves. A lot of them still think of themselves as basically getting married, having children and looking after the home. I think they feel that biology is going to be of more interest and use to them.
>
> Female biology teacher (12)

Some teachers mentioned that many girls believe that a knowledge of biology will help them into their chosen career - often a stereotypically feminine occupation such as nursing.

> Girls always want to be nurses. ... The girls think that doing biology will get them into that field.
>
> Male physics teacher (2)

A wide range of subject characteristics, believed to be associated with biology, were mentioned by some teachers as being particularly attractive to girls. They included assertions that biology is an easy subject, a descriptive subject, and that it is not technical or mathematical.

> One of the things about biology which makes it different from the other sciences is that it tends to be more descriptive. Girls appear to find this more to their satisfaction.
>
> Male biology teacher (17)

> There's less numbers and I think a lot of children are put off or frightened by numbers and formulae. Conceptually it isn't very difficult at an elementary level.
>
> Female biology teacher (19)

> It's a subject that you can learn and I suppose in a way regurgitate the knowledge. They don't necessarily have to understand it.
>
> Female biology teacher (6)

The fact that girls will be with their friends and the majority sex in a biology class was also mentioned by a few teachers.

> They know there is going to be a fair number of girls there, so there won't be the embarrassment of being in a small minority.
> Male physics teacher (4)

Besides mentioning that a girl's peer group can influence her choice of biology, parents were also mentioned by two teachers as being capable of exerting great influence.

> They are directed by parents and so on.
> Male general science teacher (13)

In decreasing order of importance, science teachers believe that girls are influenced by affective factors, social factors, subject characteristics, consideration of their future family life, consideration of their future working life, and the likely composition of the teaching group when they make their choice to study biology.

5.3.2.2 Boys reject biology

Few teachers specifically expressed opinions as to why many boys reject biology. Of those who did, three factors dominated their thoughts. Some suggested that affective factors deter boys from biology. Others that social factors and pressures deter them.

> A boy taking biology sometimes has a lot of ragging. "Oh, you going to be the only boy amongst those group of girls?"
> Female biology teacher (12)

> Males feel it's not for them, because of the traditional sex roles that people see themselves in.
> Male physics teacher (4)

A third group suggested that a scarcity of career openings in biology-related fields made biology less attractive than other subjects.

> For the boys, there aren't really careers in biology, unless they follow it to university level. And so I suppose they go for the more career orientated subjects.
> Male general science teacher (11)

5.3.2.3 Boys choose physics

The single factor mentioned by most teachers to explain why boys choose physics is its relevance for a wide range of careers.

> The boys are only too well aware that they have to find a career, and have got a long working life ahead of them. They are looking perhaps for subjects that will provide them with a positive lead into a worthwhile career.
>
> Male chemistry teacher (3)

Even if the pupils themselves are not fully aware of the vocational importance of physics, it is likely that parents will regard physics as an important subject and advise their sons to continue studying it.

> Physics tends to remain in people's minds (not so much the children's) but in people who are helping them choose, such as their parents, it tends to stick in their minds that physics is very good for engineering, mechanics, things like this. Boys are encouraged, I think, to a certain extent by parents who think of it as much better for a boy's career as opposed to a girl's.
>
> Female biology teacher (20)

The single category referred to by most teachers was that of social factors and social pressures. Tradition, society, expectations, encouragement and sex role stereotypes were all cited as influences that help to persuade boys to choose physics.

> I think it's just tradition. I think they just think that that's the natural thing to do.
>
> Female biology teacher (19)

> The way they see the rest of the world around them, boys always seem to be doing technological subjects. I think in some ways it's implied that boys are better adapted to dealing with electronic circuitry, hydrolysis, and that sort of thing. But there again, I don't think there's much of school influence. I think it's more parental, family, society.
>
> Male biology teacher (7)

> I think that's to do with expectations: That is the feelings they get that as soon as they start doing physics they take it to be a male subject.
>
> Male physics teacher (5)

Several other respondents also mentioned that physics is frequently regarded as a male subject. Teachers believe that this sex typing of physics encourages boys to choose the subject.

> Boys come into the school thinking that they are going to find physics easy. They expect to find it easy. They think of it as a boys' subject.
>
> Male physics teacher (2)

The last quote mentions two characteristics of physics, i.e. it is a male subject and it is easy. Many other teachers mentioned characteristics of physics that they believe boys find appealing. The mechanical and practical nature of physics were frequently referred to.

> They tend to choose physics because it is practical and they can get away with not sitting there and writing all the time.
> Male physics teacher (5)

> Boys, as a group, tend to like things mechanical.
> Male biology teacher (17)

Finally, a few teachers expressed the view that some boys enter physics by default, not through any positive choice on their part.

> If you are not good at physics or biology, but you have to find five subjects to fill in the timetable, then you will choose the subject where your mates are going to be also.
> Male physics teacher (2)

In summary, science teachers believe that boys choose physics mainly because of social influences (tradition, social conditioning, expectations, etc.), career considerations, the appeal of the characteristics of physics (it's practical, mechanical, etc.), and the advice of influential people, e.g. parents.

5.3.2.4 Girls reject physics

The teachers indicated that girls reject physics for three main reasons:- social influences, affective factors and characteristics of the subject that they do not like. The teachers believe that a host of social influences, e.g. tradition, sex role stereotypes, prejudice, the lack of expectations, even the lack of role models, turn girls from physics.

> I think it's a society thing, a tradition thing. Girls just do not do physical science. And the idea of course that girls don't like technical things. Role expectation. Sex expectation. We as a society don't expect girls to be interested in engines, machines, and things like that.
> Male general science teacher (13)

> They drop physics because it is not expected of them to go on to do physics.
> Male physics teacher (2)

Amongst the affective factors that turn girls from physics are their own lack of interest in the subject, and their understanding that physics is is a boys' subject.

> A general lack of interest in matters scientific, and the fact that their minds are closed to them.
> Male general science teacher (11)

> Physics is considered a boys' subject. Girls tend to not be interested. Not because they aren't that interested, but because they are encouraged not to be to a certain extent.
>
> Female biology teacher (20)

> The general impression in the outside world is that physics is (a) hard and (b) a man's subject. It's one of those things that they absorb, like women are never engineers.
>
> Male physics teacher (2)

The last quote mentions another characteristic of physics, besides its male image, that is believed to deter girls. It is the fact that physics is often described as a difficult subject. Other respondents also mentioned this point, together with other characteristics that are believed to turn girls away from physics.

> Ask most girls about physics and chemistry, they will say that they seem too hard.
>
> Male chemistry teacher (3)

> Because they don't like maths, and they've been told that in order to do physics they must be good at maths.
>
> Female biology teacher (12)

Some teachers indicated that girls drop physics because they are persuaded to do so by influential people, e.g. their parents, their friends.

> Friends and a lot of other people tell them to. Maybe dad as well.
>
> Male physics teacher (8)

Friends can not only exert direct pressure upon a girl to drop physics, but they can also exert indirect pressure through their almost certain absence from the physics class. A girl who chooses physics will probably have to accept that she will be parted from the support and company of her friends.

> Perhaps they feel that if they do it they will be swamped. They will be in a group where they will leave most of their friends. There will be a whole load of boys and they will feel rather out of it. Perhaps they feel that people will ridicule them for doing an unusual subject.
>
> Male chemistry teacher (1)

Besides the many factors discussed above that are perceived to deter girls from physics, a few teachers also pointed out that there are not many positive reasons for girls to choose physics. The pull towards physics is generally weak, compared with that of other subjects.

> One of the main reasons for choosing certain subjects is vocational/ careers. They are projecting forward and seeing themselves in careers in which physical sciences feature less importantly than biological sciences.
>
> Male physics teacher (9)

In summary, science teachers believe that girls reject physics for three main reasons:- social influences (tradition, sex role stereotypes, low expectations), affective factors (lack of interest, coupled with the masculine image of physics) and characteristics of the subject that they do not like (e.g. it's mechanical, mathematical). Subsidiary factors mentioned included negative advice from influential people, the perceived irrelevance of physics for careers, and the deterrent effect of minority sex status.

5.3.2.5 Summary

Science teachers believe that pupils' decisions to choose or reject the individual science subjects are influenced by a number of different factors. In this study the teachers indicated that the tendency for boys and girls to choose different science options (i.e. for boys to choose physical science subjects but reject biological science subjects, and for girls to choose biological science subjects but reject physical science subjects) can be largely explained in terms of five distinct clusters of factors.

1. Social factors and social pressures are recognised as powerful influences that dictate the sex appropriateness of all the science subjects. Thus such factors as tradition, social conditioning and sex stereotyped expectations dictate that girls should choose biology and drop physics, whereas boys should make the reverse decisions.

2. The characteristics associate with the different science subjects are also believed to influence the science options of both boys and girls. For instance, boys are thought to choose physics because they like the mechanical and practical aspects of the subject, and they find the subject easy. Conversely, girls are thought to reject physics because

they find the subject too difficult, and they do not like the high mathematical content. On the other hand, biology is thought to be attractive to girls because it is an easy subject, and it is not technical or mathematical.

3. Affective factors are thought to be very influential in persuading pupils to make sex appropriate science choices. The gender image of a subject is viewed as being particularly important. Thus, the teachers stated that boys choose physics because it is a boys' subject, and girls drop physics for the same reason. The reverse arguments were applied to biology. There was also considerable agreement amongst the teachers that girls choose biology because they are interested in the subject, but that they drop physics because of lack of interest. Similar arguments were not advanced to explain boys' science choices. Only two teachers suggested that boys choose physics because they are interested in the subject.

4. Pressure exerted by influential people was recognised as a factor that can affect subject choices. However, the comparative importance of this factor was seen to vary according to the sex of the pupil and the science subject under consideration. Parental advice was judged to be a relatively important factor in persuading boys to choose physics and girls to drop physics. Peer pressure was also mentioned as a factor that dissuades girls from physics. Some teachers linked this point with other deterrants. Since few girls study physics, a girl who chooses physics would not only be parted from most, if not all, of her friends, but she would also have to contend with minority sex status.

5. The relevance of a subject for future careers headed the list of individual factors believed to influence both boys' choice of physics and rejection of biology. Although girls were also thought to choose biology because of its relevance for careers and reject physics because of its irrelevance, career considerations were not thought to be so important in determining girls' science choices as they were in determining boys'

science choices.

6. A sixth important factor was mentioned, but only in association with girls' choice of biology. Girls are believed to choose biology because they think it will be useful to them in their future role as wife, housewife and mother. It is illuminating that (a) boys apparently do not conceive that biology could also be of value to them when they have a family, and (b) physics is not seen as being useful for everyday life to either sex.

5.3.3 The masculine image of science

Towards the end of the interviews, the teachers were asked why the physical sciences are often described as boys' subjects. Many of the teachers directly or indirectly implicated socialisation processes, and one specifically mentioned the potency of the attitudes held by influential people.

> It's the conditioning from our social upbringing.
> Male physics teacher (18)

> It must start pre-school, parental attitudes, junior school teachers' attitudes, toys, etc. ... I think it's parents' attitudes mainly and the peers, that's from parental attitudes of other parents.
> Male general science teacher (11)

The most commonly mentioned explanation was simply 'tradition', with over 60% of the teachers making some reference to it.

> Traditionally it is not the province of a woman.
> Male chemistry teacher (1)

> I think it's very much an ingrained idea, a traditional idea that dies very hard. There is nothing about them that makes them particularly suited for boys at all, but it is just the way that society assigns the role of men and women, even now.
> Male chemistry teacher (1)

A few teachers, after initially referring to tradition and historical reasons in general terms, then went on to detail how the educational practices of the former predominantly single sex grammar school system still influences the education offered and received in present-day mixed comprehensive schools.

> I think it is just historical.
>
> Male physics teacher (2)

> I think it might have something to do with the old grammar schools. It seemed to not be emphasized as much in girls' schools as it was in boys' schools. Now you have got comprehensives with mixed boys and girls, there's still a bit of that holding on.
>
> Female biology teacher (20)

Other teachers expanded the 'historical' explanation by pointing to the great scientists of the past who, nearly without exception, were men.

> All the great scientists you can name have been men, apart from Marie Curie, and they still are.
>
> Male physics teacher (2)

Some teachers pointed out that science in schools is dominated by males, both at the pupil level and at the teacher level.

> Simply because more boys seem to take them.
>
> Male biology teacher (7)

> Usually the teachers are males, ... and a subject where the teachers are males will be a male subject.
>
> Male physics teacher (4)

A few teachers linked the masculine image of the physical science subjects with the idea that stereotypically masculine qualities are required to study the subjects successfully.

> Traditionally boys have been better mathematicians than girls, and physics in particular involves maths. Physics often leads to interests in cars, motor bikes, which again boys traditionally are interested in.
>
> Female biology teacher (12)

The previous quote also links physics with traditionally male interests. A couple of other teachers made similar links between the physical sciences and the jobs that boys traditionally aim for.

> It's down to jobs. Boys in this area tend to be going into apprenticeships, engineering, light engineering, that sort of thing. The employers expect a physical science, preferably physics itself, not even chemistry.
>
> Male general science teacher (13)

In summary, there was general agreement amongst the science teachers that tradition and historical factors largely account for the physical sciences being described as boys' subjects. It was suggested that socialization processes and man's dominant role in science and society effectively maintain the male aura of science. No other single explanation as to why the physical sciences are viewed as male subjects was propounded by more than a very small minority of the teachers interviewed.

5.3.4 Causes of success and failure at science

5.3.4.1 Causes of academic success

The two individual variables most frequently mentioned in association with academic success in science were effort and ability. On account of the frequent reference made to these two variables, the two categories that included them - approach to work and inherent factors - were also the most frequently used categories. The approach to work category, besides including effort, also included variables such as concentration, conscientiousness, self-discipline.

> Self-discipline. Have to be able to study. Have to be methodical and systematic and tidy about what you do, if you are going to succeed academically. Got to make an effort to learn.
> Male chemistry teacher (1)

The inherent mental capacity that equips a pupil for academic studies was referred to by various terms.

> To do science you need a particular flair, a particular bent.
> Male chemistry teacher (14)

> They have got to have ability of course.
> Male physics teacher (8)

> Intelligence.
> Female physics teacher (16)

Besides possessing sufficient intelligence, pupils also need to develop certain cognitive skills if they are to succeed at science.

> There is a high requirement for them to be able to think out and apply patterns, knowledge and information to a new situation, to transfer ideas from one context to another.
> Male chemistry teacher (3)

> The ability to think logically. Looking at one thing and going on to the next step and then the next step.
> Female biology teacher (12)

Several teachers pointed out that cognitive abilities and skills alone may not guarantee success in science. Pupils also need to be interested in the subject.

> Interest is the major one, because if they are interested they motivate themselves.
>
> Female biology teacher (20)

Finally, some teachers were of the opinion that factors external to pupils can be important determinants of their chances of success at science. Two factors that were mentioned by several teachers were the family, who can give varying amounts of support and encouragement, and the teacher, who can affect a pupil's chance of success by the quality of his/her teaching.

> Primarily support and motivation and a clear plan which comes from family support as to where they are going and what they want to do after leaving school.
>
> Male physics teacher (9)

> If you can make the lessons interesting, they respond by putting in that little more effort. When it comes to difficulties, if the teacher can explain them easily in a way that they understand, that makes a difference.
>
> Male physics teacher (5)

In summary, academic success at science is principally attributed to effort and ability. Certain cognitive skills, such as logical thought, the ability to analyse and sythesise facts, are also deemed to be important contributory factors. The importance of affective factors, in particular a pupil's interest in the subject, cannot be overlooked. Finally, factors external to a pupil, especially family support and the quality of teaching received, are also believed to affect a pupil's chances of succeeding at science.

5.3.4.2 Causes of academic failure

The teachers believed that failure in science was largely due to the lack of those attributes that they had already indicated were associated with success in science. Thus some teachers said that failure was due to lack of interest and lack of motivation.

> Lower down the school, I think it's because they are not interested. They are not motivated to be interested, and they cann't be bothered to do the work.
>
> Female biology teacher (6)

A poor approach to work - lack of patience, lack of perseverance, lack of effort - were mentioned by several respondents.

> I think the English disease - laziness. Lack of effort is the single greatest cause.
>
> Male chemistry teacher (3)

Other teachers stressed the effect of lack of ability and poorly developed cognitive skills.

> They haven't mastered the techniques of learning. They haven't mastered the idea that they can learn things in patterns and remember things in patterns.
>
> Female biology teacher (10)

> The lack of basic ability, plus environmental factors that work against them, e.g. home background which doesn't value school work, peer pressures that don't value school work.
>
> Male physics teacher (4)

The last quote also mentions the adverse effect that external factors, e.g. home background, peer group pressure, can exert.

In summary, the teachers viewed failure at science to be largely due to inadequacies on the part of pupils, the most commonly mentioned being lack of ability, lack of interest and lack of motivation. A few teachers did suggest that external factors can adversely affect pupils' prospects in science. Negative parental attitudes and peer group pressure were viewed as the most serious threats to academic success.

5.3.5 Differences between girls and boys

The question "In the physical science subjects, how does the work and behaviour of girls differ from that of boys?" provoked a wide range of answers. However, a majority of the teachers included some reference to the neatness of girls' work in their replies.

> I find that their work tends to be neater, more colours used, nice underlining and diagrams tend to be neater.
>
> Male biology teacher (7)

> I have very few girls who go in for the physical sciences. Those who do are very meticulous, methodical and neat in their work. Perhaps rather over fussy sometimes. But they do well.
>
> Male general science teacher (11)

The second quote gives some suggestions as to why girls generally produce neater work. They are more conscientious and they put greater effort into their work.

> They are more conscientious. They work better. They apply themselves to what they are doing better.
>
> Male physics teacher (8)

However, some other aspects of girls' approach to their work are less conducive to success in subjects which involve significant amounts of practical work.

> I think the boys tend to be a bit more curious, fiddle with things more, want to know how they work.
>
> Female physics teacher (15)

> In chemistry, girls tend to be very much more unsure about handling the apparatus and going through a practical. They tend to take a much longer time over it, and they tend to get into a confused state much more easily.
>
> Male chemistry teacher (7)

> Sometimes the boys are a little bit better at manipulating apparatus, particularly at solving a situation where an experiment isn't working. ... But I suspect that is perhaps due to the fact that they have got a little more experience in that field, particularly if they also take physics or engineering, and if they have helped their dad tinker around with his car, etc.
>
> Male chemistry teacher (3)

Even though boys may enjoy practical work more and find it easier, the overall quality of the work of girls who choose science equals that of boys' work and often exceeds it.

> Girls cope as well with the work, certainly no worse. They get on the whole better results on average than the boys because of this difference that they have had to choose it against the sex barrier, whereas the boys tended to drift into it, so the boys will be weighted down at the bottom end.
>
> Male physics teacher (2)

The respectable standard of girls' work is not only attributed to their conscientiousness and hard work (discussed above), but also to their above average intelligence. Because the girls who study physical science are a more highly selected group than the boys, they tend to be of higher average intelligence.

> Because they have chosen to do physical sciences against the general trend, the average ability and motivation of a girl ... is higher than that for a boy.
>
> Male physics teacher (9)

Affective differences were also noted between girls and boys. Several teachers said that girls had poorer attitudes towards the physical science subjects, whilst other teachers said that girls were less interested in the physical sciences.

> Before opting, there will be a sizeable minority or possibly even a majority in a set of girls who have come into the school with the attitude that they do not like physics. They have never done physics before, but they are almost determined not to like it.
>
> Male physics teacher (4)

> There are differences in interests, which try as we can we can't always break down. Girls will persist in saying that we're not interested in this, it's a boys' subject. We take no notice of it, but they keep on saying it.
>
> Male general science teacher (11)

Some teachers spoke about differences in the behaviour of boys and girls. Opinions were divided as to whether the behaviour of girls was in fact very much better than that of boys, although most of the teachers agreed that girls were a little quieter and better behaved than boys.

> In the lessons, the people who are more active and vocal tend to be the males, but I have certainly found that there are some females who are equally active and vocal.
>
> Male physics teacher (5)

In summary, the most striking difference that teachers perceive between girls' work and boys' work is that the work of girls is neater. Nearly twice as many teachers mentioned this difference as mentioned any other single difference. Girls' attitude towards and interest in the physical science subjects was judged to be poorer than that of boys. Even so, the girls in a class are quieter and better behaved than the boys. Boys are thought to be particularly interested in and able at practical work. Girls lack confidence and manipulative experience. In spite of these disadvantages, those girls who opt to study the physical sciences produce work of a standard that equals or even exceeds the standard of boys' work. Explanations offered by the teachers included girls' higher average intelligence, and their greater conscientiousness and effort.

5.3.6 Scientists

Questions were asked in the interviews that probed three aspects of teachers' perceptions of scientists. Two questions merely sought to establish the sort of occupation conjured up by various terms used to describe scientists. Another question asked the teachers which qualities they consider to be important to a successful research scientist. Finally, the teachers were asked how scientists are commonly portrayed, i.e. how they are stereotyped.

5.3.6.1 Occupation

According to the teachers, a research scientist is a person with a higher degree, who either works at a university or in industry. The use of terms such as biologist and physicist was acceptable to the teachers and meaningful to them.

5.3.6.2 Qualities required

Nearly 90% of the teachers mentioned personality characteristics when describing qualities that a person needs to become a successful research scientist. To succeed a scientist needs to be highly motivated, very dedicated and to possess patience and perseverance. The person also needs to be very interested in science and be prepared to work exceedingly hard. As one teacher said, "Research is 99% perspiration and 1% inspiration". The ability to think laterally and originally was mentioned by other teachers as well. While some teachers noted the advantage of having an enquiring mind, others stressed the need for a logical mind. However there was general agreement that a high level of intelligence is required.

5.3.6.3 Stereotype

When the teachers were asked to describe how scientists are commonly portrayed, nearly two thirds of them said that scientists are stereotyped as eccentric/absent minded/mad professor type of people.

> Quite often as eccentric people. We get Magnus Pike and others like him, the boffin variety.
>
> Male chemistry teacher (3)

Many of the teachers described the physical appearance of scientists, especially as portrayed in TV commercials.

> White coat, row of pens in upper pocket, digital watch, calculator stuck behind the ear, roll of computer tape in a pocket: Tend to be short, spectacles, forties.
>
> Male physics teacher (2)

> People in white coats, short sighted people, bald head and glasses, male of course, that goes without saying, old before their time.
>
> Male physics teacher (4)

The sex of a scientist was indicated by a number of the respondents. They all agreed that scientists are male.

> They are always men, they are always white coated, they always have short hair, usually wear glasses. It's very rare that you see a woman scientist, and if she is, then she doesn't fulfill the sexy, fantasy type female that most men seem to have an image of. ... If you do see women on a science advert, you'd see them handing stuff to the men, like data sheets, or stood in front of the machine reading out numbers.
>
> Male biology teacher (7)

The message is clear. Men are scientists, women are assistants. Since science is regarded as a high ranking body of knowledge, the occupation of scientist carries considerable status and prestige. However, most scientists are male, and so it is men who usually benefit from the high status of the occupation.

> Scientists are thought of as an elite band who have knowledge which is in some way magic, is not open to other people.
>
> Male chemistry teacher (3)

> In school one day I was covering for a maths lesson. ... A boy said "Oh, maths is a bit of a come down Miss, isn't it, from science?" I was really astounded at that comment. ... He seemed to have science up on a pedestal, up the top somewhere.
>
> Female biology teacher (10)

Several teachers referred to the complex equipment used by scientists and the fact that they do experiments. A few teachers described scientists as being illiterate, poor communicators and uncultured. However these latter points were mentioned by a very small number of teachers.

In summary, there was much agreement among science teachers that scientists are commonly portrayed as eccentric/absent minded/mad professor type of people. In TV commercials, scientists can be recognised by their white coats and spectacles. The teachers also noted that they are always male. Another point that was mentioned by a number of teachers is the high status accorded to the occupation of scientist.

5.4 CONCLUDING COMMENTS

During the course of interviewing 23 secondary school science teachers, a number of observations were made.

1. All of the interviews took place in schools, during teachers' free periods and breaks. Most of the interviews lasted between 15-30 minutes, although a number lasted very much longer. The teachers' willingness to spend so much of their free time discussing the topics reported in this chapter indicates that they were interested in the topics.

2. All of the respondents were cooperative and congenial. They all treated the problem of 'Girls and Science' as a legitimate educational problem and gave serious and considered replies. None was hostile or dismissive of the questions put. Many of the teachers registered personal concern about the 'Girls and Science' problem.

3. Although individual teachers mentioned a number of points in reply to each question, the same points were raised by many teachers. Furthermore, the different teachers generally expressed very similar views when discussing each point, resulting in considerable uniformity of perceptions.

4. The interviews succeeded in getting science teachers to express views and opinions about a range of topics believed to be relevant to the 'Girls and Science' problem. They also proposed explanations and remedies. These discussions generated a wealth of pertinent details and additional ideas, only a small fraction of which can be included in this thesis. Appendix 5.2 contains a detailed analysis of the material that is relevant to the six main topics reported in this chapter. A short summary of the main findings relating to each topic area can be found at

the end of the six sub-sections in section 5.3.

5. One of the most striking and revealing observations to emerge from the interviews was the readiness with which science teachers spontaneously refer to physics as a boys' subject or a male subject, and the proportion of them who stereotype scientists as male.

CHAPTER 6

DEVELOPMENT OF SCALES

Page

6.0 Contents

CHAPTER 6

TABLES

FIGURES

6.1 METHODOLOGICAL CONSIDERATIONS

Many decisions have to be made during the construction of a questionnaire containing a number of scales. In this study, decisions about which variables needed to be measured and ideas about the ways in which they could be measured were guided by the hypotheses under investigation, knowledge gathered during the literature search and the views and opinions expressed by science teachers during the exploratory interviews.

The next set of decisions concerned the format of the questions and their responses. Different forms of rating scales were chosen to collect most of the data. Rating scales enable a numerical value to be assigned to some kind of judgement made by the respondents. This ensures that the scoring is objective, consistent and reliable across subjects. Furthermore, by assuming that the intervals on a rating scale are equal, the responses can then be analysed by parametric statistical techniques (see section 4.6.1). Such assumptions were made in the present study. A practical advantage of rating scales is that their structured formats take less of a research's time than do unstructured formats. Although structured formats are often more difficult and time consuming to construct, they are simple and quick to score. Closed questions, as were asked in the final questionnaires used in this study, are also relatively quick and easy for subjects to answer. This means that subjects can answer more items in a given amount of time, and also that they are less likely to become over-curious or evasive about any one item.

Having selected the types of question formats and response modes to be used in each questionnaire, individual items then had to be produced. Important aspects of scale construction, e.g. unidimensionality, linearity and equal intervals, reliability and validity were not overlooked. The approaches used in this study to establish the reliability and validity of the measuring scales are described in sections 6.3 and 6.4. Scale characteristics and their relevance are considered in Appendix 6.1, together with problems that can complicate the interpretation of rating scale responses. Besides paying due attention to the content and scale characteristics of the individual measuring devices, thought was also given to the overall design of the questionnaires formed from them. In particular, the comprehensibility, organisation and appearance of the questionnaires had to be considered. The basics of questionnaire design are discussed in Appendix 6.2.

6.2 PILOT WORK

All three main questionnaires (i.e. BIAS, STOSS and COSS) were piloted (see Appendix 4.12). The pilot stage refers to all the preliminary trying out and development that goes into questionnaire procedures to ensure that they will work as intended. The work can be of greatest help in evaluating the instructions, the questions and the response systems. Any ambiguities or lack of clarity in wording can be detected and corrected. Besides these aspects, pilot work also provides a test for the initial request and cover letter, page layout, reminders, data coding arrangements and statistical procedures. Close inspection of the results of a pilot test can show whether the questionnaire items possess the desired qualities of measurement and discriminability, and whether all the necessary information has been gathered. In addition, preliminary results will give some indication of the results to be expected from the main investigation. Careful evaluation of all the evidence procured in a pilot test should guide further decisions and actions. It might be that certain sections of the questionnaire need to be modified, that sections need to be entirely recast, or even that sections should be abandoned.

Following Youngman's advice (1978), the first pilots of the BIAS and STOSS questionnaires were personally administered to respondents. This allowed instantaneous verbal feed-back about the questionnaires to be collected. Ambiguities, omissions and muddled wording were immediately detected. Furthermore, questions which were difficult to answer or which required an inordinate amount of time to answer were detected. After modifications and improvements to the content and wording of the questionnaires had been effected, the questionnaires then had to be piloted again to check the changes that had been made. This second pilot was conducted under conditions similar to those that would exist for the main study. As with all the pilot work, the questionnaires were given to respondents as similar as possible to those in the main inquiry.

6.3 RELIABILITY OF THE MEASUREMENTS

The reliability of a measurement procedure refers to the extent to which it produces consistent results. The test-retest method of assessing the reliability of a scale is the oldest and most intuitively obvious method available. It involves administering the same instrument to the same group of people on two different occasions. Then a correlation coefficient is calculated to demonstrate similarity between the two sets of scores. The test-retest method has the advantage of requiring only one form of the scale, whereas the alternate-form method requires two equivalent forms of the scale. Also it is the only method that provides information about a scale's consistency over time. However, it has the disadvantage of being influenced by practice, memory and intervening events. To overcome these problems, internal consistency methods may be used. Cronbach's alpha, an estimate of reliability based on the average correlation among the items in a scale, provides the basic measure for determining reliability based on internal consistency (see Appendix 6.3 for the formula). Nunnally (1978) recommends that coeficient alpha should always be determined for new scales. Besides determining coefficient alpha and test-retest reliabilities for many of scales used in the present study, item analysis was also employed to increase the reliability of some scales, e.g. the Females' Social Roles

Table 6.1 Sex ratio of student and teacher samples compared

Questionnaire	Sample	N	% males	% females
BIAS	Teacher	339	67	33
	Student	21	67	33
STOSS	Teacher	164	68	32
	Student	28	64	36
COSS	Teacher	512	50	50
	Student	35	51	49

scale (see section 6.7.3.2).

Heise (1969) points out that semantic differential studies frequently focus upon mean ratings and not upon the ratings of individual respondents. In such situations, test-retest correlations between the mean scores are appropriate, and have been used by Jenkins et al. (1958) and Norman (1959). In the present study, all of the semantic differential results are discussed as scale means and thus the reliability of the scales is best determined by test-retest correlations between the mean scale values. Many of the other scales used in the study also focus upon the mean ratings assigned to items rather than upon the ratings of individual respondents. Test-retest correlations between mean ratings is again an appropriate measure of their reliability and stability.

Reliability coefficients were obtained for all the scales that appeared in the final form of the three main questionnaires. Values for the individual scales are presented at the end of the sections describing their development, i.e. sections 6.6.1 to 6.9.1.

PGCE students from UDEs were enlisted to help with the reliability studies. This decision was taken in view of the difficulties encountered in contacting large numbers of science teachers in secondary schools. Furthermore, it was believed that practising teachers would be very reluctant to complete a questionnaire twice to allow reliability measures to be estimated by the test-retest method. Although the student teacher samples were younger than the teacher samples contacted in the main surveys, they were specialised to teach a similar range of subjects and their sex ratios were very similar to those of the teacher samples (see Table 6.1). Furthermore, even though the students' teaching experience was more limited, they had all spent several months in schools on teaching practice. Thus they were considered to be adequately qualified to substitute for teachers. In addition, the BIAS sample of student teachers were particularly well qualified to mark chemistry write-ups since 76% of the sample were training to be chemistry teachers. Finally, it should be remembered that the student sample could very well have been picked up in a school sample a few months later.

The time interval between the two administrations of each of the three questionnaires (BIAS, STOSS, COSS) was approximately one month. This is generally considered to be a satisfactory interval (Henerson et al., 1978). However, the small sample sizes and their homogeneity reduced the likelihood of obtaining high reliability coefficients.

The acceptability of a particular reliability coefficient depends upon the proposed use of the scale. Much lower reliabilities can be tolerated for scales designed to discriminate between groups than can be tolerated for tests designed to discriminate between individuals. Within groups, the error scores in the observed scores of individuals tend to cancel each other out when the observed scores are averaged across individuals. The degree of cancelling out increases with group size, and so lower reliability coefficients can be tolerated with larger groups. Values of 0.60 to 0.70 are usually regarded as acceptable (Open University, 1973), regardless of the type of reliability coefficient calculated (Henerson et al., 1978). Guilford (1954) suggested that values as low as 0.50 may be justifiable for research purposes, if the alternative is to discard the scale. In the present study, reliability coefficients of 0.60 were deemed to be acceptable.

Reliability is important because it is a necessary condition for validity. The validity of a measure can never exceed the square root of its reliability (Lord & Novick, 1968). Thus, reliability limits validity. However, high reliability does not of itself guarantee validity. The validity of a scale must be established independently of its reliability (Bohrnstedt, 1970). If a scale has no validity, then it is useless. Generally, a valid measure is also a reliable measure.

6.4 VALIDITY OF THE MEASUREMENTS

The validity of a measuring instrument refers to the extent to which the instrument measures what it purports to measure (Tuckman, 1978). Thus a valid instrument measures the characteristic that it is designed to measure and not some other characteristic. Strictly speaking, it is not a measuring instrument that is validated, but rather some use to which the instrument is to be put (Nunnally, 1978). The validity of a measuring instrument is dependent upon context.

Since the particular application of an instrument determines its validity, there are therefore four different types of validity.

(a) Content validity refers to the representativeness of the sample of questions included in the instrument.

(b) Concurrent validity means that the instrument produces results that correlate well with those from other measures of the same construct*.

(c) Predictive validity refers to the instrument's ability to predict some future behaviour.

Concurrent validity and predictive validity are often classified as criterion-related validity since they both evaluate an instrument against some criterion that is assumed to be valid.

(d) Construct validity refers to the extent to which the test appears to conform to predictions about it made from theory.

Hence studies of construct validity are concerned with validating the theory underlying the instrument. This is done by investigating whether or not the instrument confirms or denies the hypotheses predicted from a theory that is based on the constructs.

In the present research, it was not possible to access the concurrent validity of any of the scales used as no pre-existing scales

* The word 'construct' is taken to have a wide meaning, as proposed by Henerson et al. (1978). "It is a catch-all term used to refer to the skills, attitude, or ability that an instrument is intended to measure" (p.135).

could be found that measured the identical constructs to those under investigation in this study. Predictive validity was not applicable either since none of the measurements were designed for predictive purposes.

The development of the scales, which is detailed in sections 6.6 to 6.9, ensured content validity. A search of the literature and discussion with science teachers provided a detailed specification of the constructs under investigation. It is unlikely that any important aspects of the constructs have been overlooked. Thorough piloting of the scales during their developmental stage allowed redundant items to be discarded and extra items to be added for facets which were inadequately sampled. All of these procedures helped to ensure that the final scales adequately sampled the different facets of each construct investigated.

The construct validity of an instrument can be defended in a number of ways (Cronbach & Meehl, 1955; Henerson et al., 1978). The internal structure of an instrument provides some evidence. A scale which has good internal consistency has a high degree of homogeneity and is obviously measuring a unitary construct. The reliability of a scale can also be used to support construct validity. If a construct is hypothesized to be a stable characteristic of individuals, then scores should remain stable over time. High test-retest correlations cannot prove that the measure of a construct is valid, but low correlations would challenge its validity. Thus the test-retest and alpha coefficients presented in the present study not only indicate the reliability of the scales, but also support their construct validity.

The circumstances of the administration of a scale can affect the scale's validity. Rushing respondents or pressurizing them to respond in a particular way can invalidate results. Details regarding the administration of the scales in the present study can be found in section 4.5. As far as is known, all respondents were allowed adequate time to complete the questionnaire, there was no communication between respondents until after they had filled the questionnaires, and neither the format of the questionnaires nor the explanations and directions contained therein should have influenced the replies of the respondents. Thus, if it is accepted that the instruments are logically related to the constructs that they are designed to measure, then their validity was not violated during administration. This argument for construct validity is particularly applicable to the marking exercise. Comments received from respondents indicated that they were unaware of the true nature of the investigation. Thus they could not have consciously injected sex bias into the results.

Close inspection of the results obtained from a scale can throw light upon the scale's validity. If the scale seems to be measuring constructs other than those intended, e.g. reading comprehension rather than scientific competence, then the construct validity of the scale is highly suspect. Application of this construct validation technique to the scales used in the present study does not cast doubt upon the validity of any scale. Other construct validation techniques were also used, but since each technique applied to only a few scales, they are reported in Appendix 6.4.

6.5 SCALE DEVELOPMENT DETAILS: INTRODUCTION

The following five sections describe the development and/or use of all the scales employed in the investigation. The first four sections consider those scales that were designed to gauge teachers' views regarding the masculinity of the science subjects, to measure the extent to which they hold sex stereotyped ideas regarding their pupils and the science subjects, and to indicate whether they display sex bias in their expectations and judgements. The first section covers the development of scales to investigate the sex typing of science, the second section covers scales designed to reveal sex stereotyping, the third section describes the development of attribution scales, and the fourth section describes the measures that were developed to investigate teachers' expectations and judgements. The small number of purely fact gathering questionnaires that were used in the investigation are referred to in the fifth section.

The development and use of every scale used in the study is summarised in Appendix 4.12, Table A4.12/4. Full details regarding the administration of the scales and the characteristics of the samples to whom they were administered can be extracted from Table A4.12/2 in Appendix 4.12, and Table A4.11/1 in Appendix 4.11. Table 4.2 may also prove helpful. It locates descriptions of scales, their development, the results obtained, and ensuing discussions.

6.6 SCALE DEVELOPMENT DETAILS (A) SEX TYPING OF SCIENCE

6.6.1 School Subject Characteristics

A measure was required that would allow different groups of teachers' views about the different school subjects to be compared. The semantic differential was chosen as the measuring instrument. Although the method is more usually used to investigate connotative meaning (as in section 6.6.2), it is also well suited to collecting data of a straightforward descriptive nature. Some of the technique's advantages are outlined below.

(a) Semantic differential scales are comparatively easy for a researcher to produce, administer and code.

(b) They are also easy for respondents to complete.

(c) The technique is very economical, since one set of scales can be used to measure perceptions of many concepts.

(d) Because (c) applied, semantic differentials provide a generalized approach and allow direct comparisons to be made between the perceptions of different concepts.

(e) The procedure has been widely used and has demonstrated reliability and validity.

During the development of the semantic differential scales, several steps were taken to ensure their validity. Special attention was paid to the bipolarity of the adjective pairs, their relevance to the school subjects under investigation, and the length of the rating scales. In addition, the choice of school subjects was also given careful consideration.

6.6.1.1 First pilot

A set of 14 adjective pairs was compiled. Some were extracted from lists supplied by Osgood et al. (1957), others were derived from previous research specifically concerned with the image of subjects (Weinreich-Haste, 1979, 1981), and yet others were suggested by opinions expressed by science teachers during the exploratory interviews (see section 5.3.1).

Four school subjects were chosen for initial investigation. The science subjects were represented by physics, probably the most masculine science subject (see section 2.1.1), and biology, the least masculine science subjects. Geography was chosen as a sex neutral subject (Murphy, 1980; Ormerod et al., 1979) and French as a feminine/arts subject (Ormerod et al., 1979; Weinreich-Hasts, 1979).

In the first pilot, the four school subjects were presented in a fixed sequence to sample C. Each school subject was followed by all 14 scales. The order and polarity of the scales was kept the same for all four school subjects. However, the polarity direction of the scales was alternated to ensure that the teachers' responses would swing from end to end of the scales (e.g. 'arts-science' but 'masculine-feminine'). This format is recommended to prevent respondents from developing response sets which could reduce the sensitivity of the measurements (Osgood et al., 1957).

Each semantic differential scale consisted of 7 rating points, bounded by the contrasting adjectives. The scale positions were not labelled with appropriate adverbs, although adverbial quantifiers were supplied in the instructions. A study by Wells and Smith (1960) indicated that respondents are better able to differentiate between the different rating positions when adverbial labels are provided. Also, labelled scales are easier to understand, which results in greater cooperation from respondents. The adverbial labels provided in the instructions were 'extremely', 'quite' and 'slightly', since these labels define rating positions that are approximately equal distances apart (Heise, 1970).

6.6.1.2 Second pilot

The replies to the first pilot provided useful information about the clarity of the instructions, the relevance of the scales, and the response rate to be expected from participating schools. However, the representativeness of the results was suspect as the respondents taught in independent schools. Therefore the study was repeated with a sample of teachers from comprehensive schools (sample F).

The results from the first and second pilots were compared for differences. The t test for independent samples indicated that the independent teachers gave significantly different ratings ($p < 0.05$) to the state teachers on only 3 of the 56 (4x14) rating scales. Since this ratio is barely greater than would be expected by chance at the 5% probability level, the amalgamation of the two sets of results seemed justified.

6.6.1.3 Third pilot

A third pilot was conducted in order to gather still more data upon which to base final decisions regarding the school subjects to be investigated and the semantic differential scales to be used. The third pilot was similar to the first two pilots, but with the following alterations.

(a) 3 of the 14 semantic differential scales were removed and replaced by 2 new rating scales. Thus each school subject in the third pilot was rated on 13 scales.

(b) The semantic differential scales consisted of 6 rating positions instead of 7. Some researchers (Weinreich-Haste, personal communication) maintain that the use of a 6-point scale forces respondents to examine their perceptions and to record a definite answer. The third pilot investigated this claim.

(c) The four school subjects investigated were physics and biology (reasons as before), plus English literature and home economics. The latter two subjects were included as alternative feminine subjects (Murphy, 1980; Ormerod et al., 1979).

(d) The order in which the school subjects were presented was varied. Half of the respondents rated physics, biology, English literature and home economics, whilst the other half rated English literature, home economics, physics and biology.

Preliminary analysis of the data obtained in the third pilot from sample H, investigated the effect of a school subject's position in the booklet upon the ratings awarded to it. The mean ratings received by a subject when it appeared in two different positions in the booklet were compared using a series of t tests for independent samples. Of the 52 comparisons made, 6 were significantly different ($p < 0.05$). This proportion is higher than would be expected by chance (2.6/52). Therefore it would seem that the ratings awarded to the school subjects were not entirely independent of the order in which the subjects were presented. This finding does not accord with results reported by Osgood et al. (1957) and Sommer (1965). They found that concepts were judged independently of the conceptual context in which they were embedded. Heise, writing in 1970, reassured researchers that the order in which concepts are presented in a test booklet is immaterial since anchoring or order effects do not operate. Nevertheless, on the basis of evidence that contrast effects were operating in the third pilot, a decision was made to vary the position of each school subject in the final from booklet.

Decisions regarding the format and selection of the semantic differential scales to be included in the final questionnaire, were made on the basis of the results obained from all three pilot studies.

6.6.1.4 Choice of rating scales

The process of rejecting unsatisfactory semantic differential scales and of retaining useful adjective pairs was guided by the following considerations.

(a) The range of the ratings awarded to the different school subjects on each adjective pair was inspected. Adjective pairs which produced mean ratings that differentiated between science/masculine and arts/feminine subjects by more than 2 points were judged to be satisfactory.

(b) The dispersion of the ratings awarded on each adjective pair was considered. Adjective pairs that produced average standard deviations of 1.5 and greater were deemed to be unsatisfactory.

(c) Adjective pairs whose bipolarity was suspect (see section 6.6.1.5) were judged to be unsatisfactory.

(d) Adjective pairs that produced significant correlations with the masculine-feminine semantic differential were judged to be satisfactory.

(e) Adjective pairs that produced significant correlations with the science-arts semantic differential were judged to be satisfactory.

(f) Adjective pairs that received significantly different ratings from different groups of respondents (e.g. male/female teachers, science/non-science teachers) were judged to be satisfactory.

The final criterion refers to pilot work concerned with teachers' perceptions of pupils' preferences for subject characteristics. This pilot work was conducted at the same time as the pilot studies on school subject characteristics.

(g) Adjective pairs describing characteristics which were perceived to

Table 6.2 Percentage of teachers who endorsed the bipolarity of the COSS adjective pairs (N = 49)

Adjective pair	Percentage
Theoretical - practical	93.9
Numerical - verbal	75.5
Science - arts	95.9
Logical - intuitive	77.6
Feminine - masculine	100
Factual - opinionative	75.5
Creative - routine	79.6
Simple - complex	98.0
Important - unimportant	100

be preferred to a significantly different extent by boys and girls were deemed to be satisfactory.

Guided by the above criteria, 9 adjective pairs were selected to be included in the final questionnaire.

6.6.1.5 Bipolarity of adjective pairs

The bipolarity of the 13 adjective pairs used in the third pilot was checked by sample J. Each teacher was given a total of 21 adjective pairs arranged in two lists. The second list consisted of the jumbled antonyms of the adjectives in the first list. They were asked to pair the adjectives. It was made clear that the jumbled adjectives could each be used once, more than once or not at all in forming the pairs. Agreement amongst the respondents' replies ranged from 29% to 100%. Eleven of the 13 adjective pairs elicited agreement from more than half of the respondents. Poor agreement amongst respondents was taken as an indication that an adjective pair should be discarded, since its meaning, when applied to school subjects, was obviously ambiguous.

Once the adjective pairs to be used in the final questionnaire had been picked, their bipolarity was checked again. A different sample of teachers (sample L) was used for this exercise. The procedure outlined above was repeated using the nine adjective pairs. All of the adjective pairs were correctly identified as being opposite in meaning by more than three quarters of the respondents. The detailed results are recorded in Table 6.2.

6.6.1.6 Number of rating positions

To compare the ratings awarded to physics and biology using the 6 and 7 point scales, the mean ratings obtained on the 6-point scale were converted to equivalent 7-point values using the following formula

$$X = 1 + \left(\frac{Y - 1}{6 - 1} \times 6\right)$$

where Y = mean rating on 6-point scale
X = equivalent rating on 7-point scale.

Table 6.3 Mean ratings awarded on 7-point (N=112) and transformed 6-point (N=57) semantic differential rating scales

	Physics		Biology	
S.D. scale	7-point	6-point	7-point	6-point
Practical-theoretical	3.73	3.90	2.92	3.59
Factual-opinionative	1.66	1.84	2.38	2.28
Arts-science	6.76	6.81	6.13	6.10
Animate-inanimate	5.02	4.87	2.38	2.62
Objective-subjective	2.03	1.95	2.68	2.60
Masculine-feminine	3.11	3.27	4.30	4.49
Numerical-verbal	2.20	2.41	4.44	4.28
Structured-open-ended	2.47	2.60	3.31	3.32
Creative-routine	4.61	5.04	4.03	4.73
Intuitive-logical	6.00	6.05	4.73	5.08
Simple-complex	5.73	5.78	5.10	5.04

Note The scoring system awarded 1 to the rating position nearest the left hand adjective in a pair, and 7 to the rating position nearest the right hand adjective. 4 represents a neutral rating.

The converted 6-point ratings, together with the 7-point ratings, for each characteristic are recorded in Table 6.3. Inspection of the table shows that the use of a 6-point scale may have forced the teachers to express opinions about the characteristics of each subject under investigation, but that the mean ratings were generally less extreme than those obtained with a 7-point scale. Only one characteristic of physics, its routine nature, was clearly rated more extremely on the 6-point scale than on the 7-point scale. For biology, 3 of the 11 characteristics were clearly rated more extremely on the 6-point scale. The adjectives are 'feminine', 'routine' and 'logical'.

It is customary to use 7-point scales in semantic differential research (Heise, 1969). The practice was first recommended by Osgood et al. (1957) on the basis of findings from early methodological research. In view of the fact that the 6-point scale used in the third pilot study generally produced mean ratings that were very similar to the mean ratings obtained from a 7-point scale, it was decided to use a 7-point scale for the final questionnaire. This decision was endorsed by the observation that respondents were more reluctant to complete 6-point scales than 7-point scales. In fact, several of the teachers in the third pilot spoilt their returns by attempting to give neutral responses.

6.6.1.7 Choice of school subjects

Subjects were required which would span the arts/science spectrum and the masculine/feminine spectrum. The following considerations guided the final selection.

(a) Include subjects previously identified as very masculine or very feminine (Weinreich-Haste, 1979, 1981).

(b) Include subjects previously identified as being boys' subjects or girls' subjects (Hutt, 1979).

(c) Include subjects preferred by boys and subjects preferred by girls (Ormerod et al., 1979).

Table 6.4 Coefficient alpha reliabilities and average test-retest reliabilities on the School Subject Characteristics scale

School subject	Coefficient alpha	Average r	No. of significant correlations
Physics	0.67	0.60	8/9
Chemistry	0.58	0.47	8/9
Biology	0.18	0.59	9/9

Table 6.5 Mean score reliabilities on the School Subject Characteristics scale

School subject	Mean score correlation (r)
Physics	0.98
Chemistry	0.98
Biology	0.97
Maths	0.99
French	0.98
History	0.97
Home Economics	0.97
Woodwork	0.99

(d) Include subjects with more than 55000 entries from one sex at GCE O level or CSE.

(e) Include subjects in which boys or girls obtain a distinctly higher pass rate in external examinations.

(f) Include subjects trialed by joint CSE/GCE consortia in connection with 16+ feasibility and development studies.

(g) Include at least one subject from each broad subject area mentioned in Statistics of Education summaries.

(h) Include those subjects most frequently mentioned in the exploratory interviews as being typical arts or science subjects.

(i) Include at least one subject from the different subject areas that were to be investigated in this study for their perceived importance (see section 6.7.4).

Guided by the above nine criteria, particularly the CSE and GCE O level entries, 12 school subjects were chosen for detailed study. They were physics, chemistry, biology, maths, history, geography, English (language), French, art, home economics, woodwork and technical drawing.

6.6.1.8 Reliability

Sample ST(COSS) (N=35) were used to investigate the reliability of the final scale when applied to a selection of school subjects. All three science subjects were included in the study and so was maths. Alpha coefficients were calculated for the three science subjects (see Table 6.4). Test-retest reliabilities were calculated from the replies of a sub-sample of ST(COSS) (N=20). The average correlation between the ratings awarded to each science subject on the two occasions is recorded in Table 6.4, together with the proportion of the individual item correlations for each subject that were significant at the 5% level. One-tailed tests were used since positive correlations were predicted. Replies from the sub-sample of ST(COSS) were also used to calculate correlations between the mean scores awarded to each subject on the first and second administrations. The r values obtained for each school subject included in the study are recorded in Table 6.5. The values are very high indicating that the mean scores display reliability and stability. Despite a couple of low reliability coefficients in Table 6.4, the School Subject Characteristics scale was judged to be sufficiently reliable for its intended use. This judgement was largely based upon the high mean score correlations produced by the scale.

6.6.2 Masculinity Index

The purpose of the Masculinity Index was to measure teachers' feelings about the masculinity of each of the three main science subjects. By measuring feelings it was hoped to get beyond expressed opinions or rationalized thoughts to more fundamental beliefs and attitudes. An index was chosen, rather than a single question, because indices generally produce more reliable and valid measures of an attitude (Henerson et al., 1978).

In view of the abstractness and obscurity of the concept being measured, i.e. the gender of a science subject, it was decided that the semantic differential would be an appropriate measuring instrument. The semantic differential is generally regarded as a good tool for measuring affect towards an attitude object. It is particularly useful in situations where respondents have emotional reactions to a topic, but are unlikely to be able or willing to express them freely as statements. A semantic differential score represents the respondent's general feelings or impressions about the attitude object.

Preliminary stages in the development of the Masculinity Index involved the choice of adjective pairs, clarification of their gender connotations, and verification of their bipolarity.

6.6.2.1 Adjective pairs

It was necessary to pick a number of adjective pairs which could be applied to school subjects, which would have no obvious denotative meaning when applied to school subjects, and which could be shown to possess gender connotations.

Initially the book 'The measurement of meaning' by Osgood et al. (1957) was searched for suitable adjective pairs and a number were extracted. The list was eventually reduced to seven adjective pairs. Three of them were retained in their original form and the other four were altered slightly.

6.6.2.1.1 Gender connotations

The first study to establish whether one adjective of each pair was associated with masculinity and the other with femininity was conducted with sample B. For each adjective pair, respondents were asked to write 'M' beside the adjective that they felt was most closely related to 'masculine', and 'F' beside the one most closely related to 'feminine'. The results (which are reproduced in Appendix 6.5) were quite promising. A clear demarkation between the gender connotations of the two adjectives in each pair emerged. However, since the sample was a group of Women's Institute members, it was recognised that the study needed to be repeated using a more appropriate sample, i.e. a group of teachers. Before administering the scale to the teachers, one more adjective pair was added.

In the second study (sample D), a number of respondents returned unuseable replies. They either refused to complete the scale, signified that they associated both adjectives of a pair with the same gender, signified that each adjective was both 'masculine' and 'feminine', or signified that the adjectives were neutral. The full results appear in Appendix 6.5.

To overcome the teachers' reluctance to clearly sex type the adjectives, it was decided to ask them to rate each adjective individually on a 7-point rating scale running from 'extremely masculine'

Table 6.6 Percentage of teachers who correctly paired the Masculinity Index adjectives (N = 34)

Adjective pair	Percentage
Active - passive	88.2
Hard - soft	100
Weak - powerful	76.5
Tender - tough	67.7
Cold - warm	97.1
Intimate - remote	76.5
Light - heavy	97.1

to 'extremely feminine'. The mid-point was clearly marked 'neutral'. This new scale was completed by sample E. The results (which are recorded in Appendix 6.5) demonstrated very clearly the teachers' proclivity to rate each adjective as neutral.

In the next study a 6-point rating scale was used. Only the extreme ends were labelled 'very masculine' and 'very feminine'. This scale seemed to be acceptable to sample J for few refused to complete it. In addition, it produced unambiguous results (see Appendix 6.5). If the respondents had genuinely viewed each adjective as neutral, one would expect roughly half the respondents to rate it on the masculine side and half on the feminine side. In fact, for nearly every adjective, a large majority of respondents chose either a masculine rating or a feminine rating. In addition, one adjective of each pair was rated masculine and the other feminine. Thus, the gender connotations of each adjective pair were established.

6.6.2.1.2 Bipolarity

The bipolarity of the adjective pairs was checked by a sample of secondary teachers (J). The Masculinity Index adjective pairs were combined with a number of distractor adjective pairs to form a list of 21 pairs. Each teacher was given the adjectives arranged in two lists. The second list consisted of the jumbled antonyms of the adjectives in the first list. They were asked to pair the adjectives. It was made clear that the jumbles adjectives could each be used once, more than once or not at all in forming the pairs. The results are shown in Table 6.6.

All of the adjectives were correctly paired by more than three quarters of the respondents, with the exception of the tender-tough pair. However, it was decided to retain this pair as it is highly sex typed (see Appendix 6.5, Table A6.5/5).

6.6.2.2 Scale construction

McKennell's (1970) attitude scaling procedure was used to construct

Table 6.7 Factor loadings for the Masculinity Index items

(A) PHYSICS

Item	Principal factors (I)	(II)
1. Active - passive	-0.11	-0.45
2. Hard - soft	0.74	-0.17
3. Powerful - weak	0.64	-0.42
4. Tough - tender	0.77	-0.23
5. Cold - warm	0.70	0.43
6. Remote - intimate	0.67	0.49
7. Heavy - light	0.83	-0.03
8. High - low	0.27	-0.29

(B) BIOLOGY

Item	Principal factors (I)	(II)
1. Active - passive	-0.07	-0.34
2. Hard - soft	0.86	-0.37
3. Powerful - weak	0.15	-0.54
4. Tough - tender	0.69	-0.27
5. Cold - warm	0.56	0.60
6. Remote - intimate	0.52	0.27
7. Heavy - light	0.46	-0.08
8. High - low	-0.42	-0.30

the scale. This method is described in greater detail in Appendix 6.3. Although the item pool (see section 6.6.2.1 for its selection) was presented to samples A&A* and I&I* along with the other scales of the STOSS questionnaire, their replies were not used for the Masculinity Index pilot study. It was considered that the samples were too small and homogeneous for the task. Instead, samples C and F were used to pilot the items for the Masculinity Index. They were piloting the COSS questionnaire anyway, which consisted of semantic differential scales, and so two scales (for physics and biology) were just lengthened slightly to include the Masculinity Index items. (Appendix 7.4 considers in detail a number of points that justify this decision) Preliminary analysis of the results revealed that there were no significant differences between the replies to the Masculinity Index items from the two samples (C and F) and so the replies from the two samples were pooled. This produced a combined sample of 112 respondents.

6.6.2.2.1 Dimensionality

Initially cluster analysis (see Appendix 6.6) was used to investigate the dimensionality of the collection of items. For both physics and biology, the items fell into two clusters. Following McKennell's advice, the composition of these clusters was checked using factor analysis (see Appendix 6.7). It was found that the first principal factor for each subject accounted for a substantial proportion of the variance in the matrix from which it had been derived, 45.3% in the case of physics and 33.1% for biology. The factor loadings are shown in Table 6.7. If items 1, 3 and 8 are disregarded, the same strong general factor runs through each matrix. On account of these general factors obtained by factor analysis, the pool of items were regarded as being unidimensional.

6.6.2.2.2 Item selection

Before proceeding to alpha-scaling, the homogeneity and validity of

Table 6.8 Estimated reliability of the Masculinity Index after discarding successive items

(A) Physics

Item No.	20	18	16	19
n	4	3	2	-
$\bar{r}_{ij}^{*}$	0.48	0.49	0.49	0.54
$\bar{r}_{ij}$	0.50	0.52	0.48	-
Alpha	0.80	0.76	0.65	-

(B) Biology

Item No.	19	20	18	16
n	4	3	2	-
$\bar{r}_{ij}^{*}$	0.30	0.38	0.39	0.42
$\bar{r}_{ij}$	0.37	0.45	0.66	-
Alpha	0.70	0.71	0.80	-

Key

n = number of items remaining after items on the left have been discarded

$\bar{r}_{ij}^{*}$ = average correlation of each item with the other three items

$\bar{r}_{ij}$ = average intercorrelation of the n items

Alpha = reliability value for the n items

Item 16 hard - soft

18 tought - tender

19 cold - warm

20 remote - intimate

the eight items was considered. Four items were rejected on the basis of their low factor loadings on the first principal factor in factor analysis, their low correlation with the response to a direct measure of the gender of the science subject (a 'masculine-feminine' semantic differential item), or their negative correlation with a few of the other items in the masculinity index.

Coefficient alpha was calculated for the remaining four items. The value of alpha provides an estimate of the reliability of a scale composed of those four items. The effect upon the reliability of the scale of shortening it by removing items was also investigated. See Table 6.8, and Appendix 6.3 for a fuller discussion of the procedure. On the basis of the information contained in Table 6.8, it was decided to retain all four items.

6.6.2.3 Reliability

The reliability studies conducted on the final scale with sample ST(STOSS) yielded alpha coefficients of 0.70, 0.72 and 0.69 when applied to physics, chemistry and biology respectively. Test-retest reliability was also calculated for a sub-sample of ST(STOSS) (N=12). The correlation between the sum of the ratings given to physics on the first and second administrations was 0.70. The comparable figures for chemistry and biology were lower.

6.6.3 Characteristics of Science

A scale was required that would chart science teachers' perceptions of the characteristics of the science subjects in some detail, and facilitate the identification of any characteristics that are particularly closely associated with science's masculine image.

6.6.3.1 First pilot

A number of adjectives that are commonly applied to the science subjects were extracted from the literature. This produced a list of 44 words and phrases. The teachers were asked to indicate the degree to which they believe that their principal teaching subject possesses each of the listed characteristics. Four response options were supplied, ranging from 'extremely' to 'not at all'.

The replies received from sample A&A* suggested that the meaning of two items was not clear. In addition, four items were completely failing to differentiate between the physical and biological sciences. Thus six items were removed from the list and another five items were added.

6.6.3.2 Second pilot

The list of words and phrases presented to sample I&I* in the second pilot was essentially the same as that used in the first pilot. This enabled more replied to be received from teachers of each science subject. Replies were eventually received from 34 physical science teachers and 23 biology teachers.

The verbal rating scale was converted to a numerical scale of 1-4, and then the mean ratings awarded to the physical science subjects and to biology were compared for each item (see Appendix 6.8 for mean ratings). Those items that failed to differentiate sufficiently between the two subject areas were discarded. An arbitary difference value of at least 0.5 was chosen to determine those items which were retained. This procedure produced 20 items. Upon closer inspection, some of the items were judged to have overlapping meanings, and so a further five items were rejected.

The original intention had been to ask each teacher to rate all three science subjects on all the items in the final questionnaire. However, it was decided that this might be too tedious for the respondents. Hence, the two subjects which are generally accepted to be most dissimilar (physics and biology) were chosen.

6.6.3.3 Reliability

The reliability studies conducted on the final scale with sample

ST(STOSS) yielded alpha coefficients of 0.52 and 0.63 for physics and biology respectively. The low values of these coefficients suggest that the scale covers a range of weakly related characteristics of science, most probably is not unidimensional, and thus has weak internal consistency. Since the scale was designed to produce mean scores for a school subject, and not to register differences between the ratings awarded by individual respondents, attention should be focused upon the magnitude and stability of the mean scores. The profile of the mean scores should remain stable over time. To test this, the mean scores obtained on the first and second administrations were correlated. A value for r of 0.96 was obtained for physics and of 0.93 for biology.

6.6.4 Opinions

Since an opinion is an expressed attitudes or belief (Open University, 1975), it was thought appropriate to construct some form of attitude scale to gauge teachers' opinions regarding the reasons for the physical science subjects' masculine image. A scale with the item format developed by Wilson and Patterson (1968) recommended itself as being particularly appropriate for the task.

Wilson and Patterson (1968) proposed that the evaluative propositional statements of traditional attitude scales could be replaced by single words or phrases. The subject would respond by indicating those items he or she favours or believes in. The response to each item would thus reflect the subject's gut reaction to the idea or concept presented and would be little mediated by the problem-solving thought evoked by items in sentence form. Wilson and Patterson (1968, 1973) suggested that by focusing upon attitude content alone, the influences of context (grammatical confusion, ambiguity, task conflict, cognitive processes, acquiescence, social desirability) could be reduced to a minimum. Similar arguments have been posed by Kerlinger (1972) to support the construction of attitude scales from referent items rather than statement items.

Wilson and Patterson (1968, 1973) used brief nondirectional catch-phrases to construct a Conservatism Scale. They showed it to have high reliability and validity. In spite of some criticism (Pedhazur, 1978; Ray & Pratt, 1979; Schneider, 1973), the scale has been widely used (Ray & Pratt, 1979) and the format generally accepted.

The advantages of employing a similar format to construct a scale to tap opinions in this study were that such a scale would be very quick and easy to complete. Thus a wider coverage of ideas would be possible.

6.6.4.1 Pilot

Items from which to construct a scale were obtained by scanning the literature for suitable ideas and by referring back to the replies received in the exploratory interviews to the two questions:

(a) Some educational researchers suggest that science has a masculine image. What do they mean by this?

(b) Why are the physical sciences often described as boys' subjects?

These two activities produced 43 items, which were expressed as single words or short phrases.

Two similar scales were formed. First, respondents were asked to indicate those factors which, in the eyes of the general public, contribute to the masculine image of the physical science subjects. Then they were presented with the same scale again, but asked to show their own personal opinions about the causes of the masculine image of the physical science subjects. Opinions regarding the explanatory value of each item were recorded using three response alternatives: 'yes' '?' and 'no'. These are the same response options as those used by Wilson and Patterson (1968). '?' allowed respondents to register misunderstanding, indifference or neutrality, although respondents were discouraged from choosing this response option.

Preceding the two opinion scales were four short, direct questions.

The first two asked whether the respondent thought that the general public regard the physical and biological sciences as masculine, feminine or neutral subjects. The other two questions made the same enquiry, but the respondent had to indicate his or her own perceptions of the gender of the two science areas.

The first pilot, which was administered to sample C, showed that respondents were prepared to express opinions about the explanatory value of each item. On some items, the respondents' replies produced nearly 100% agreement, whilst on others their opinions were split nearly evenly. The '?' category was rarely used. Respondents were overwhelmingly prepared to express an opinion one way or the other.

Careful consideration of the pilot results indicated that the data were not in a form appropriate for making comparisons and interrelationships with results from other scales. In the interest of keeping all questionnaires as brief as possible, it was decided that the Opinion scales would not be included in the final STOSS questionnaire. Instead, in view of the satisfactory content validity of the scales, they were used immediately, without further alteration or piloting. Copies of the scales can be seen in Appendix 6.9.

6.6.5 Scientist Stereotypes

Sadker and Sadker (1982, p.245) define a stereotype as "An uncritical or oversimplified belief regarding the characteristics of a particular group; this belief is based on the assumption that because members of the group share one characteristic, they are similar in many others." This definition implies that certain characteristics are associated with scientists simply because they are scientists. Furthermore, if scientists are stereotyped as men, then by extension of the argument, additional characteristics will be associated with scientists, not necessarily because they are scientists, but simply because scientist is equated with man.

The stereotyped image of scientists has been investigated by a number of researchers. Scientists are viewed as being remote, scruffily dressed, eccentric, socially isolated, hard working, very intelligent, logical and masculine (Ashton & Meredith, 1969; Beardslee & O'Dowd, 1961; Bendig & Hountras, 1958; Hills & Shallis, 1975; Hudson, 1967; Mead & Métraux, 1957; Selmes, 1969). Other studies investigating scientists' personality traits have identified that successful scientists frequently have a non-verbal bias of intelligence, low sociability, and are independent, sef-sufficient, confident, persistent and more concerned with things than people (Cattell & Drevdahl, 1955; Roe, 1952; Terman, 1955). Thus some of the descriptions that appear in the stereotype of a scientist accurately reflect the personality of many scientists. However other terms, e.g. eccentric, remote, are more tenuously linked with reality. They constitute a stereotype - a fixed impression which conforms very little with the facts it purports to represent.

Work on people's stereotypes of scientists has been conducted using various types of respondents. Most studies have used schoolchildren or students, but some have used the general population, including scientists and people working in science-related fields. A range of research methods have been used. According to Hesselbart (1978), there are five different formats that are frequently used in stereotype research. They are (a) adjective checklists, (b) percentage measures, (c) semantic differential scales, (d) Likert scales, and (e) group comparisons.

The semantic differential was chosen to investigate stereotypes in the present study. This choice was made largely through consideration for the respondents. As respondents would already be familiar with the technique, having previously completed the Masculinity Index, they would be able to answer speedily. Other advantages of the semantic differential technique are listed in section 6.6.1. Having decided upon the research method, decisions also had to be made regarding the content and format of the semantic differential scales, and the concepts to be rated.

6.6.5.1 Choice of scientists

It was felt that if science teachers were asked to indicate their perceptions of 'a scientist', they would find the term too vague. Most science teachers are graduates, and most graduate teachers have specialized in only one of the science subjects (findings from pilot work and educational statistics). Therefore, science teachers should have well differentiated views about scientists from the different science disciplines. It is possible that they even differentiate between the different specialities of a discipline, e.g. zoologist, botanist, geneticist, ecologist. To check whether the terms biologist and physicist would be satisfactory, science teachers were asked in the exploratory interviews "Is the term biologist or physicist meaningful to you?" The majority of respondents answered 'Yes' (11/12), and only one respondent expressed any reservations. This result suggested that biologist and physicist were appropriate terms to use in subsequent studies. Physicist and biologist were chosen because they occupy positions at opposite ends of the gender dimension (see literature review). Chemist was omitted in order to keep the questionnaire short.

6.6.5.2 Format of rating scales

The number of rating positions along each semantic differential scale was seven. The choice of this number was strongly influenced by the findings and arguments presented in section 6.6.1.6.

Considerable thought was given to the actual format of the semantic differential scales. It was felt that if the usual format (as used in sections 6.6.1 and 6.6.2) was used, respondents would complain that they were being asked to make generalizations of an unreasonably broad and blanket-type nature. In addition, the inclusion of a gender scale would be problematical. The use of the adjective pair 'masculine-feminine' was rejected since the intention was to focus upon the perceived biological sex of scientists, not their gender related characteristics.

The desired meaning could be conveyed by a 'male-female' scale, but such a scale would effectively have only two rating positions. To overcome these problems, it was decided to use semantic differential rating scales to ask respondents about the probability of scientists possessing certain characteristics.

The use of probability judgements to measure stereotypes is consistent with psychological theory. A stereotype has already been defined as an uncritical or oversimplified belief regarding the characteristics of a particular group. The term belief refers to the degree of acceptance of a proposition regarding the characteristics of an object or event. Fishbein has written extensively about beliefs (Anderson & Fishbein, 1965; Fishbein & Raven, 1962) and his ideas can be summarize thus, "Beliefs are assumptions about the probability that an object exists, that it possesses certain characteristics, or that it is related in certain ways to other objects" (Open University, 1976, Block 10, p.11). Fishbein advocates that beliefs be measured by asking respondents to rate relational statements on a series of bipolar probability scales (e.g. probable-improbable, likely-unlikely) of the semantic differential form (Fishbein & Raven, 1962). In the present study probable-improbable rating scales were used to measure the teachers' beliefs about the characteristics possessed by scientists.

6.6.5.3 Choice of characteristics

6.6.5.3.1 First pilot

In the first pilot (sample A&A*), respondents were asked to rate the probability of the same eight characteristics being possessed by a physicist and by a biologist. These characteristics included one ascribed characteristic (male) which is usually linked with scientists, several characteristics reportedly possessed by successful scientists, e.g. self-sufficient, and several characteristics that are reported to be stereotypically associated with scientists, e.g. highly regarded.

The mean ratings indicated that the teachers believed that a physicist was more likely than a biologist to possess all of the characteristics, with the exception of self-suffi iency and dedication. These two aforementioned characteristics were also comparatively ineffective at differentiating between a physicist and a biologist. Since it was decided to reject the three characteristics with least discriminatory power, 'self-sufficient', 'dedicated' and 'highly regarded' were removed from the list of characteristics. They were replaced by another three characteristics which were chosen not only for their reported contribution to the stereotype of scientist, but also for their contribution to the male sex stereotype.

6.6.5.3.2 Second pilot

The revised list of characteristics was completed by sample I&I*. It was found that all of the characteristics satisfactorily differentiated between a physicist and a biologist, and that the stereotypes operated in the predicted directions. Therefore, no further changes were made to the scale.

6.6.5.4 Reliability

The reliability studies conducted on the final scale with sample ST(STOSS) yielded alpha coefficients of 0.65 and 0.57 for physicist and biologist respectively. The mean scores obtained on two separate admininstrations of the scale correlated highly. The value of r for physicist was 0.98 and for biologist it was 0.94. Thus the scale reliably produced similar profiles of mean scores from one administration to the next.

6.7 SCALE DEVELOPMENT DETAILS (B) SEX STEREOTYPING

6.7.1 Written Work of Girls and Boys

In the first administration of this scale to sample A&A* the teachers were asked two questions:

(a) Would you say that you can generally distinguish between the written work of boys and girls?

(b) Can you briefly indicate any features that you consider to be typical of the written work of girls and boys.

The first question was a forced choice question and teachers could either reply 'yes' or 'no'. The second question was open-ended. The wording was deliberately neutral, so that the respondents' answers would not be directed or biased by the question's frame of reference or the questioner's preconceived ideas.

The replies indicated that the teachers were willing to answer the questions, and that the wording of the questions was clear. Therefore the format (see Appendix 6.10) remained unchanged in subsequent administrations of the scale.

6.7.2 Preference for Subject Characteristics

6.7.2.1 Pilots

The semantic differential scales, that were developed to ascertain teachers' impressions of the characteristics of different school subjects (see section 6.6.1), were also used to determine teachers' perceptions of pupils' preferences regarding subject characteristics. Teachers were asked to use these scales to indicate the extent to which they believe pupils prefer the different subject characteristics. By presenting the scales twice, teachers' beliefs about the preferences of a typical 14-year-old girl and a typical 14-year-old boy could be obtained.

A detailed description of the development of the semantic differential scales can be found in section 6.6.1. The Preference for Subject Characteristics scale was piloted, along with the School Subject Characteristics scales, using samples C, F and H. Decisions regarding the content and length of the Preference for Subject Characteristics scale were largely based on the results from the School Subject

Characteristics scales. The same semantic differential items were used in both scales to allow direct comparisons to be made between the results.

6.7.2.2 Reliability

The mean scores obtained from sample ST(COSS) on two administrations of the scale correlated highly. The value of r for girls' preferences was 0.95 and for boys' preferences it was 0.85.

6.7.3 Females' Social Roles

Over the past fifteen years a number of scales have been developed that measure attitudes towards women, females' social roles, feminism, women's liberation, etc. Whilst reviewing the literature over 20 such scales were located. However, they were all developed and have mainly been used in the United States of America. It is doubtful whether conditions are sufficiently similar on either side of the Atlantic to justify the use of an American scale in England, without first trialling it. Moreover, there would be no information regarding the reliability and validity of such a scale when used with English subjects. For these reasons an extensive search was undertaken to find a British scale. One which had been developed by Pauline Slade was eventually located (Slade & Jenner, 1978). More recently Parry (1983) has adapted the American Attitudes towards Women Scale (Spence & Helmreich, 1972) and produced a British version. Unfortunately this scale did not appear until after the present work was completed.

Although Slade and Jenner's (1978) Attitude to Female Role questionnaire met the criterion of being a British scale, insufficient was known about its psychometric properties. Decisions taken whilst developing the scale had been made on the basis of replies from a very small sample (N=20). The scale had only ever been administered by one worker and to a limited range of subjects, all of whom were female. Although information about the reliability and validity of the scale was available, no study had been made of the scale's factor structure. Before the scale could be used in the present study it was obviously necessary to find out more about the actual structure of the scale, and to check that its use with men presented no unforeseen problems. Also it was decided to produce a shortened version of the questionnaire. A short scale can be nearly as reliable as a much longer scale, and is quicker to complete. Moreover, a long scale seemed inappropriate in the present study, as it would unnecessarily alert respondents to unstated aspects of the investigation, e.g. teachers' beliefs about sex role stereotyping.

6.7.3.1 First pilot

All 25 items of Slade & Jenner's (1978) Attitude to Female Role questionnaire were given to a mixed sex group of teachers (sample K), similar to those to whom the final scale would be administered, i.e. science teachers. The items had 4-point Likert response formats ranging from 'strongly disagree' to 'strongly agree'. An undecided category was not included. Items were scored 0, 1, 3 and 4. This scoring system, which is designed to force an expression of attitude, was used by Slade and Jenner. The total score was obtained by addition. The full results appear in Appendix 6.11.

The sample size (N=45) was considered to be rather small for the planned analyses, especially factor analysis (see Appendix 6.7). Difficulties had been encountered in achieving even the obtained sample size, so it was decided that it would be more productive to analyse the gathered data using a variety of techniques, a multi-method approach, rather than attempt to substantially increase the size of the sample.

Before proceeding to select items for a short scale, the reliability and factor composition of the whole scale was first determined.

Table 6.9 Details of possible short versions of the Attitude to Female Role scale

No. of items in scale	Item numbers	Estimated reliability
5	10,11,12,16,24	0.87
10	2,10,11,12,13,15,16,19,21,24	0.91
15	2,9,10,11,12,13,14,15, 16,19,20,21,22,24,25	0.91

Reliability was assessed using coefficient alpha (see Appendix 6.3), an internal consistency measure. A value of 0.89 indicated acceptable internal reliability. The scale was factor analysed (see Appendix 6.11) to determine its dimensionality. Seven factors with eigenvalues greater than one were extracted. The first unrotated factor accounted for 31.9% of the variance, and no other factor accounted for more than 9.7%. This indication of a single factor structure compares favourably with values obtained for other attitude to women scales, e.g. the FEM scale (Singleton & Christiansen, 1977; Smith et al., 1975).

The first factor loadings obtained through factor analysis provided a convenient criterion for the selection of scale items. Retention of those items with the highest loadings ensured that the resulting scale would comprise a relatively unidimensional measure of the attitude. Item-whole correlations were also calculated (Tuckman, 1978). This item analysis procedure helps to identify those items that show the greatest amount of agreement with the total score. The selection of those items with high correlations produces a scale with good internal consistency and consequently good reliability. The third item analysis technique used was one described by Edwards (1957). The 27% of the subjects with the highest total scores and the 27% of the subjects with the lowest total scores were identified as high and low scoring groups (Engelhart, 1972). Then the difference between the mean score of the high and low groups on each item was calculated. Items giving large differences were selected as they obviously differentiated well between the two criterion groups. The results obtained from this analysis, together with those from the two techniques described previously, are presented in Appendix 6.11.

The selection of items to produce a shortened scale was guided by the results appearing in Appendix 6.11. Table 6.9 shows the best items to pick for a 5, 10 or 15 item scale. Coefficient alpha reliability estimates (Appendix 6.3) for each scale are also included in the table.

The alpha values indicate that any of the proposed short version scales would display good reliability. However, inspection of the content of the items selected for the 5 and 10 item scales casts doubt upon the usefulness of such scales. All of the items in the proposed 5 item scale state or imply that women should be subordinate to men. Furthermore, "a woman's place is in the home". Even if the scale is lengthened to 10 items, the same sentiments still predominate. Only by lengthening the scale to 15 items does a little more variety enter the scale. Since a 5 or possibly 10 item scale was considered to be an appropriate length for the present study, it was reluctantly decided that a short version of the Attitude to Female Role questionnaire would not be suitable. The items in such a scale would be too repetitive.

6.7.3.2 Second pilot

Once the Slade and Jenner (1978) Attitude to Female Role questionnaire had been rejected, there was no alternative but to investigate the feasibility of using an American scale. Two of the most widely used scales, the AWS (Attitudes toward Women Scale) and the FEM (Attitudes toward Feminism) scales, were considered (Smith et al., 1975; Spence & Helmreich, 1972). Both scales have been used by a number of researchers and have been given to a wide range of subjects. Furthermore, the factor structure, reliability and validity of the scales are well established. The FEM scale has also been produced in shortened (5 and 10 item) versions (Singleton & Christiansen, 1977). After careful consideration the AWS was rejected on the basis that it was the older scale and the content of some of the items appeared stilted and dated. Furthermore, although a short (25 item) version of the scale has been produced (Spence et al., 1973), it is still too long for the present study.

Before the FEM scale could be used, it was obvious that the wording of some of the items needed to be altered in order to make the meaning clearer. The assistance of six judges was enlisted to identify those

Table 6.10 Details of possible short versions of the FEM scale

No. of items in scale	Item numbers	Estimated reliability	Correlation with 20-item scale
5	2,4,5,7,18	0.86	0.89
10	2,3,4,5,7,10, 14,15,17,18	0.86	0.94

items which needed to be altered, and to gather suggestions for their improvement. As a result of this consultation, three items were altered slightly and one item was totally rewritten.

The modified 20 item scale was administered using the less emotive title of 'Females' Social Roles' to a mixed sex group of science teachers (sample L). The items had 5-point Likert response formats ranging from 'strongly agree' to 'strongly disagree'. Items were scored 1 to 5; 1 being the most traditional response, 5 the most liberal. The total score was obtained by addition. The full results appear in Appendix 6.11. As in the first pilot, the sample size was disappointingly small. Again, a decision was taken to analyse the data using a multi-technique approach.

First the reliability and factor composition of the whole scale was determined. Reliability was assessed using coefficient alpha. A value of 0.87 indicated acceptable internal realiability. The scale was factor analysed (see Appendix 6.7) to determine its dimensionality. Six factors with eigenvalues greater than one were extracted. The first unrotated factor accounted for 30.2% of the variance, and no other factor accounted for more than 10.7%. This indication of a single factor structure compares favourably with values obtained for the FEM scale by Smith et al.(1975), and Singleton and Christiansen (1977).

The criteria used to select items for a short version of the scale included first factor loadings, item-whole correlations and mean score differences between high and low scorers. The values obtained for these three different analyses are reported in Appendix 6.11. The values for each analysis were ordered on the basis of decreasing magnitude and then those items with the highest values were selected. This procedure resulted in a unidimensional scale, i.e. one that was internally consistent and hence reliable. Table 6.10 shows the best items to pick for a 5 or 10 item scale. Both factor loadings and item-whole correlations indicated this selection. The same selection was also indicated by mean score differences, with the exception of one item for both the 5 and 10 item scales.

Coefficient alpha reliability estimates (Appendix 6.3) are recorded in Table 6.10. They indicate that either of the proposed short version scales would display good reliability. Correlations between scores on the shortened versions and the original 20 item FEM scale are also included in the table. The values show that scores on both of the short scales correlate highly with scores on the full set of items. Consideration of the alpha and correlation values, together with the need for a short attitude scale, suggested that the 5 item scale would meet requirements and adequately discriminate between teachers holding traditional and more liberal beliefs about woman's role in society.

6.7.3.3 Reliability

The reliability study conducted on the final scale with sample ST(STOSS) yielded an alpha coefficient of 0.86. Test-retest reliability was also calculated for a sub-sample of ST(STOSS) (N=12). The correlation between the total score for all 5 items achieved on two separate occasions was 0.90.

6.7.4 Importance of Subjects

A scale was required that would tap teachers' opinions about the importance of different school subjects. It was decided to concentrate upon those subjects that often become optional when pupils reach 14 or thereabouts. The number of subjects currently taught in secondary schools is very great, so it was first necessary to group the subjects into a more manageable number of categories. A search of the literature provided ideas for possible categories (HMI, 1979; Hutt, 1979). Eventually, after consulting a number of secondary teachers, it was decided to use seven groups of subjects. English, maths, religious education and physical education were omitted from the list, because they are compulsory subjects for most pupils in most schools (HMI, 1977).

6.7.4.1 First pilot

Sample H, consisting of teachers of all subjects, participated in the first pilot study. The respondents were asked about their beliefs concerning the importance of each of the seven subject areas to the general education of pupils. The question was asked twice. The first time with reference to the general education of boys, and the second time with reference to the general education of girls. The two questions appeared immediately beneath each other. The teachers indicated their assessment of the importance of each subject area to each sex on a 4-point scale, ranging from 'very important' to 'not at all important'.

6.7.4.2 Second pilot

The wording of the question used in the first pilot was rather abstract and could have evoked a theoretical answer from the teachers, such as would apply in an ideal world. Therefore, a second pilot was conducted using a more precise question. Teachers were asked how important they thought CSE/O level qualifications in each of the seven subject areas would be to pupils in their future lives. Again the question appeared twice. The first time referring to girls and the second time to boys. The same 4-point rating scale was used. The replies indicated that the revised question was satisfactory, and so no further changes were made.

6.7.4.3 Reliability

In section 6.3, attention was drawn to the fact that a valid measure is generally also a reliable measure. Since this scale is clearly content valid for measuring the perceived importance of a wide range of optional school subjects, the high content validity should be accompanied by acceptable reliability. Although the reliability of the scale when administered to teachers was not established, a reliability study of the same scale when administered to pupils was conducted. A sample of 206 secondary school pupils completed the scale on two occasions, separated by a time interval of approximately one month. The average correlation (r) between the ratings awarded to each subject area on the two occasions was 0.60. The correlation between the mean scores awarded to each subject area on the first and second administrations was 0.98. This value is very high indicating that the mean scores display reliability and stability. Although the work reported above does not constitute a proper reliability study, since it involved a very different sample of respondents, nevertheless it does provide additional support for the assertion that the scale, as used in the present research, was reliable.

6.8 SCALE DEVELOPMENT DETAILS (C) ATTRIBUTION PATTERNS

6.8.1 Reasons for Success/Failure at Science

6.8.1.1 First pilot

Both theoretical and empirical considerations guided the construction of the initial scale designed to investigate sex differences in the attribution patterns of teachers. The scale was composed of the four causal factors customarily included in attribution studies (Weiner et al., 1972), together with another four factors that emerged in the exploratory interviews (see section 5.3.4). The instructions requested teachers to rank the five factors that they believe contribute most to pupils' success or failure in science. Each scale was produced in two formats: one referring to boys, the other to girls. Furthermore, teachers were asked to complete the scale with particular reference to their own teaching subject. Thus the scale encompassed a number of variables - academic performance (success or failure), pupil sex (boy or girl), principal teaching subject of teacher (biology, chemistry, physics), and teacher sex (male or female).

Inspection of the replies immediately revealed that many of the teachers (sample A&A*) had failed to follow the ranking instructions.

This may have happened because the instructions were not clear, or because the teachers found the task inappropriate or too difficult, and so chose to indicate their responses in an alternate form.

A detailed analysis of the useable replies (N=23) showed that two of the factors which had been provided to explain success - task simplicity and luck - had hardly been used by any of the teachers. Likewise, two of the failure factors - illogical mind and bad luck - had been grossly under-utilised.

Comparing the attributions made for boys' and girls' success in science, the patterns were very similar. The attribution patterns for failure were also similar for boys and girls, with the exception of attributions to lack of motivation. The teachers indicated that they believe that girls' failure is more likely to be due to lack of motivation than is boys' failure.

In view of the scale's inability to detect differences in attribution patterns between girls and boys, and bearing in mind that the wording of the instructions was unsatisfactory, a decision was taken to discard the first Success/Failure scale.

6.8.1.2 Second pilot

A search of the literature was conducted to establish the content of attribution scales that have been used in previous work with teachers. It was discovered that most investigators used scales that had originally been developed from free response attributions given by pupils to explain their success or failure at academic tasks (Bar-Tal & Guttmann, 1981; Lorenz, 1982). Only one study, by Cooper and Burger (1980), was found which reported categories derived from the free response attributions of practising teachers. Seventeen categories were listed. However, after comparing them with those used by other workers, Cooper and Burger suggested that the list could be shortened to 12 categories. These 12 categories were used to construct the second pilot scale for the present investigation. One additional category, that of motivation, was also included since the first pilot study had indicated its potential as a sex differentiating factor.

Teachers were asked to rate the importance of each of the 13 factors on a 4-point scale ranging from 'very important' to 'not at all important'. Each teacher rated the factors with regard to girls' success and failure or boys' success and failure in their teaching subject area. Thus, as in the first pilot, the scale encompassed the variables of academic performance, pupil sex, teacher sex and principal teaching subject of teacher.

Analysis of the results from the second pilot, conducted with sample I&I*, indicated that the teachers were failing to differentiate between the three categories of attitude, motivation and interest when describing the causes of boys' success and failure. Therefore, it was decided to omit the motivation category from the final form of the scale (since motivation was an extra factor to Cooper and Burger's suggested scale), but otherwise to leave the scale unaltered.

6.8.1.3 Reliability

The reliability studies conducted on the final scales with sample ST(STOSS) gave an alpha coefficient of 0.61 for the success scale and a coefficient of 0.68 for the failure scale. The mean scores obtained on the first and second administrations correlated highly. The value of r for the success scale was 0.95 and for the failure scale it was 0.93. Thus the scales gave very similar profiles of mean scores, the scores of interest, from one administration to the next.

6.8.2 Reasons for Choosing/Dropping Science

6.8.2.1 First pilot

In the absence of any previous studies into teachers' perceptions of the reasons why pupils opt to drop or continue with a school subject, the pilot scales were derived from ideas expressed in the exploratory interviews (see section 5.3.2). Three pertinent questions had been asked:

(a) Why do more girls than boys study biological subjects?

(b) Why do more boys than girls study physical science subjects?

(c) Why do most girls drop physical science when choosing their subject options at 13+ or thereabouts?

The replies to these direct questions produced a number of categories. The most commonly employed categories and/or the most commonly mentioned factor within each category were taken to construct the scales. There were 11 variables in each scale.

Teachers were asked how frequently they believe each reason applies when pupils make their subject options at 13+ or thereabouts. They indicated their responses on a 5-point scale ranging from 'never applies' to 'always applies'. Each scale was produced in two formats: one referring to boys, the other to girls. Furthermore, teachers were asked to complete the scale with particular reference to their own teaching subject. Thus the scale encompassed a number of variables - subject choice (choose or drop), pupil sex (boy or girl), principal teaching subject of teacher (biology, chemistry, physics), and teacher's sex (male or female).

The replies received from sample A&A* showed that the instructions had been followed correctly, and that the reasons that appeared in the scales were meaningful. Suggestions from some teachers led to the addition of one extra reason to the scale, and the rewording of two reasons so as to increase their accuracy.

6.8.2.2 Second pilot

In the second pilot the two scales, now composed of 12 reasons, were given to sample I&I*. As in the first pilot, the teachers rated each reason on a 5-point scale. Individual teachers rated the reasons with regard to girls' choosing and dropping or boys' choosing and dropping their principal teaching subject. Thus, as in the first pilot, the scale included the variables of subject choice, pupil sex, teacher sex and teaching subject of teacher.

Analysis of the results showed that interest, or lack of it, was considered to be a very common reason determining subject choice. When all the reasons were rank ordered, interest appeared at the top of the list or in the second position, regardless of science subject, pupil sex or subject choice considered. Since one of the main objectives of the scales was to establish any differences between the perceived motivation of boys and girls when choosing or dropping science subjects, it was decided to omit interest from the final scale. Social pressure was also omitted because it received very similar ratings to tradition, indicating that the teachers were failing to differentiate adequately between the two factors.

6.8.2.3 Reliability

The reliability studies conducted on the final scales with sample ST(STOSS) gave an alpha coefficient of 0.59 for the Reasons for Choosing scale and a coefficient of 0.76 for the Reasons for Dropping scale. The mean scores obtained on the first and second administrations correlated highly. The r values were 0.98 and 0.94 for the choosing and dropping scales respectively. Thus the scales gave very similar profiles of mean scores, the scores of interest, from one administration to the next.

6.9 SCALE DEVELOPMENT DETAILS (D) TEACHER EXPECTATION AND TEACHER JUDGEMENT

6.9.1 Marking Exercise

The purpose of the marking exercise was to investigate whether science teachers display sex bias in their expectations and judgements by asking them to mark samples of pupils' written work, which were sometimes presented as boys' work and sometimes as girls' work. The final experimental design is described in greater detail in section 4.3.2.

A number of tasks were associated with the initial development of the marking exercise. First, samples of experimental write-ups and homework essays had to be collected. Names had to be found for the fictitious pupils. The different aspects of the work that the teachers would assess had to be chosen, and it was also considered desirable to establish the overall standard of each work sample when it was presented in a neutral context.

6.9.1.1 Obtaining the samples of work

The experimental write-ups came from a set which had originally been collected in association with another research project, and was subsequently made available for this work. The write-ups, which were of experiments on distillation, had been produced by 12-year-old boys and girls shortly after they started secondary science. The pupils had been introduced to the ideas of physical change and change of state. Then after a brief discussion about the process of distillation, they were presented with the practical problem of producing a sample of distilled water from tap water using only simple apparatus. After selecting their apparatus and carrying out the distillation, the pupils had to write up their experiment. The whole set of write-ups were inspected closely and then three specimens which seemed to be representative of good, average and poor standards, were chosen for inclusion in the marking exercise.

The homework essays were specially produced for this research. A class of 12-year-old pupils at a local comprehensive were asked to write an essay entitled "What I think about science and scientists". The essays were sorted on the basis of the standard of writing, and the nature of the attitudes expressed. The three specimens which were selected, were judged to be interesting to read, and to contain both positive and negative statements about science and scientists.

6.9.1.2 Choosing the names

Names were required to attach to the work samples to denote the sex of the pupil who produced the work. However, several workers have shown that names are linked with stereotypes (see section 2.5.2.2.1). These stereotypes and associated expectancies can influence our perception of a person, and can influence how other information acquired about the person is processed and interpreted. It has been shown that teachers expect higher academic standards of pupils with popular names (Nelson, 1977), and even award them higher marks (Harari & McDavid, 1973). In view of the known biasing influence that names can exert, a decision was taken to use popular names of comparable favourability for the marking exercise, in the belief that such an action was most likely to minimise the influence of the names' stereotypes upon the marks awarded to the samples.

6.9.1.2.1 First study

To pick a set of potentially popular names for further investigation, use was made of the relationship between familiarity and popularity. Familiar first names are generally preferred to unfamiliar names (Colman et al., 1981). The initial study was ambitious in its attempt to link a first name with a surname. Seven surnames were picked on the basis of their frequency in the local telephone directory. Smith was the most common surname (29 columns), whilst Taylor and Williams were the least common surnames used in the investigation (10 columns each). Seven boys' names and seven girls' names were picked in a fairly arbitary manner. However, they were assumed to be of comparable popularity and to be non-class denoting.

Each surname was linked with a boys' and a girls' first name, and respondents (sample B) were asked to indicate how well they liked each of the names on a 5-point scale.

The results (which appear in a condensed form in Appendix 6.12) showed that none of the names were unanimously liked or disliked. In addition, the ratings given to most of the boy/girl pairs were quite dissimilar. For some pairs the boy version was preferred and for others the girl version was preferred.

The conclusions drawn from this study were that a more rational basis for choosing first names was required, and that the use of first names alone might well produce greater consensus from the raters than the use of first names with surnames. In addition, it was recognised that the sample, a group of Women's Institute members, was inappropriate and that a sample of teachers should be used for the next study.

6.9.1.2.2 Second study

Two of the surnames used in the previous study were both linked with two popular boys' names and two popular girls' names, giving a total of eight first name/surname combinations. To these were added the two most popular boys' and girls' names registered in 1950 and in 1975 (Dunkling, 1977). The complete list was then rated by secondary school teachers (sample D).

The findings (see Appendix 6.12) showed that first name and surname combinations were largely unsatisfactory, since there was insufficient agreement over their popularity. There was greater consensus about the popularity of the single first names. Hence it was decided to investigate a greater range of first names, and to include first names that have been popular over longer periods of time.

6.9.1.2.3 Third study

Eight girls' names and eight boys' names were chosen on the basis of their consistent popularity over the last century (Dunkling, 1977). Because the popularity of boys' names tends to be remarkably stable, the boys' names chosen actually headed the popularity lists for 1950 and 1975.

Secondary school teachers rated the list of 16 first names and indicated how well they liked each name (sample E). The results (see

Appendix 6.12) showed that three names - Sarah, David, Michael - were generally viewed favourably. To confirm this finding and to detect additional popular names, it was decided to repeat the study using an extended list of names.

6.9.1.2.4 Fourth study

The names chosen for this study were taken from a variety of sources: the more promising names from the previous study, the most popular names registered for newborn boys and girls in 1975 (Dunkling, 1977), and the most popular names which appeared in the birth announcement column of The Times during 1980 (Brown, 1981). The list consisted of a total of 24 names. A sample of secondary school teachers (sample J) indicated their personal preference for each of the names.

Condensed results are recorded in Appendix 6.12. They indicate that some of the names were definitely more popular than others, and that there was a high degree of agreement over the popularity of some names. Emma was clearly the most popular girls' name, followed by Claire. Rebecca was chosen as the third most popular girls' name on the basis of its comparatively low rating in the 'dislike' category. Matthew and Paul were the most popular boys' names on the basis of their ratings in the 'like' category. Mark was chosen as the third most popular boys' name because it had the lowest rating in the 'dislike' category.

Having established the three most popular girls' and boys' names in the list, they next had to be paired for use in the marking exercise. This was guided by the distribution of their ratings in the three categories of 'like', 'uncertain' and 'dislike'. Names with most similar distributions were paired together. Thus Emma was paired with Matthew, Claire with Paul, and Rebecca with Mark. These pairings also reflect the relative popularity of the names. It was decided to attach Emma/Matthew to the work sample pair of highest standard, Claire/Paul to the samples of average standard, and Rebecca/Mark to the work samples which were of below average standard.

Table 6.11 Factors used by science teachers to assess written work (N=16)

Category	No. of mentions
Scientific accuracy	
Scientific accuracy	15
Content	6
Appropriate use of technical terms	5
Relevance	3
Accuracy in observation	2
Organisation of ideas	
Logical coherence	11
Acceptable account format	6
Standard of diagram	
Clear diagram	9
Appropriateness of diagram	2
Clarity of explanation	
Clarity of explanation	9
Understanding of principles	
Understanding of principles	8
Awareness of limitations of experiment	3
Originality	3
Deductions drawn	2
Grammar and spelling	
Spelling	8
Grammatical accuracy	6
Presentation	
Presentation	6
Neatness	6
Legibility	2
Completeness of work	
Completeness of work	6
Thoroughness of explanation	2
Characteristics of child	
Type of child	5
Child's ability	5
Child's background	2
Assistance given	2
Age of child	2
Effort involved	
Effort involved	5
Length of work	2

6.9.1.3 Reproducing the samples

The three experimental write-ups and the three homework essays were paired together in a fairly arbitary manner, since the three homework essays were considered to contain comparable attitudes towards science and to be not too dissimilar in overall standard. The three sample pairs were then accurately reproduced in three different handwriting styles, all of which were easily legible. The pupil's fictitious name was written at the top of each sample of work, and all original errors were faithfully copied.

6.9.1.4 Variables used in marking

In order that science teachers would be asked to rate the samples of work in the marking exercise on realistic variables, it was thought advisable first to check with science teachers how they usually mark written work. A number of science teachers (sample A&A*) were asked which factors they usually take into account when assessing written work. Because the question was open-ended, a variety of replies were received. These replies were coded using the method suggested by Henerson et al. (1978). The first answer was inspected for the factors it contained. Each different point was recorded and a tally stroke (/) was placed beside each to indicate that the point had been made in one answer. The rest of the replies were then read and new factors were recorded as described above. When points were encountered that had already been mentioned in a previous answer, another tally stroke was made to signify the additional mention. Table 6.11 contains all the different factors which were mentioned by two or more respondents. The factors have been arranged into broad categories, and ordered according to the number of mentions each factor received. The broad categories have been labelled with a heading that attempts to summarize the contents of the category. These categories, together with another one taken from the Written Work of Girls and Boys scale (section 6.7.1), formed the factors on which science teachers were asked to rate the work sample pairs in the pilot stage of the marking exercise.

Besides supplying the rating variables for the marking exercise, this study also evoked ideas about possible approaches to the judgement of the overall standard of the work samples. Some respondents indicated, especially for the poorer work samples, that they would award higher marks than the work objectively merited. Most of the teachers also supplied reasons to support this action. As a result of these comments, the pilot questionnaire asked teachers to indicate both the marks that each work sample merited and also the marks that they would actually award it.

6.9.1.5 Establishing the standard of the work samples

The literature review indicates that the marks awarded to a piece of work can be influenced by the labels attached to the work and by the context in which it appears (see section 2.5.2.2.1). Thus it was thought to be prudent to establish the perceived standards of the experimental write-ups and of the homework essays separately and with no indication of the pupils' sex. These standards would assist in allowing any subsequent biases which were detected to be ascribed to context or label effects. Furthermore, it would allow cases of biased marking to be categorized as over-marking or under-marking.

The three experimental write-ups, with no names attached, were rated on a 3-point scale by a number of science teachers (sample A&A*). 71% of the teachers agreed that the first piece of work entitled 'Distillation of water' was of average standard, 79% rated the second piece of work, 'A dislitted evention' (sic), as being of below average standard, and 85% judged the third piece of work, 'My apparatus for separating pure water from tap water' to be of an above average standard. Thus the teachers clearly considered that the three pieces of work were of differing standards, and there was general agreement over the standard of

Table 6.12 Standard (and percentage agreement) of unnamed work samples

Write-up/essay number	1	2	3
Standard of write-up	Average (71%)	Below average (79%)	Above average (85%)
Standard of essay	Average (67%)	Above average (60%)	Average (58%)
Names added to sample pair	Claire/ Paul	Rebecca/ Mark	Emma/ Matthew

each piece of work.

The three homework essays, with no names attached, were rated on a similar 3-point scale by another sample of science teachers (sample K). Agreement over the standard of each essay was not so good as for the write-ups, but for each essay one scale position received a majority vote. 67% of the teachers judged the first essay to be of average standard, 60% rated the second essay to be of above average standard, and 58% thought that the third essay was of average standard.

Table 6.12 summarizes the standard of each experimental write-up and each essay, when they were rated without names attached. The percentages in brackets show the percentage of teachers who agreed with the judgement. The table also indicates the way in which the work samples were combined together for use in the marking exercise, and the boy's and girl's name attached to each sample pair.

The intention had been to select three essays of average standard, since they were supposedly written by pupils of average ability. Table 6.12 indicates that the standard of the second essay was higher than that of the other two essays. However, it was considered that this slight discrepancy did not constitute a serious departure from the intended situation. Subsequent pilot work supported this contention (see Appendix 6.13).

6.9.1.6 Pilot work

The marking exercise was piloted on secondary school science teachers (sample I). They were asked to mark three pairs of write-ups and essays, each pair having been written by a different pupil. Besides giving an overall mark to the work, the teachers also rated a number of cognitive and attitudinal factors.

Close inspection of the completed questionnaires suggested that the overall task was acceptable, the instructions were adequate, the wording was unambiguous, and the rating variables were meaningful.

Table 6.13 Dimensions underlying the ratings awarded by teachers in the pilot marking exercise

1. Grammar and spelling Standard of work Organisation of ideas Clarity of explanation Conciseness Completeness of work
2. Aptitude for science Scientific accuracy Understanding of principles Effort involved
3. O level suitability CSE suitability
4. Attitude towards science Interest in science
5. Mark merited Mark given
6. Neatness Standard of diagram

Some of the simpler statistical analyses that were planned for use with the final data were tried out on the pilot data. Unfortunately, because the sample size was rather small, the results had to be regarded as being of a tentative nature. In spite of this, some trends and relationships did emerge, e.g. a relationship between teaching chemistry and marking the work samples more severely.

Particular attention was paid to the relationships between the different rating variables in order to gauge their suitability for retention. As a first step, the mean values given to the three sample pairs for each variable were compared, and correlations were calculated between the ratings given to each variable summed across all three sample pairs. Next the variables were grouped to denote underlying dimensions. The primary method used was a form of cluster analysis, although the usefulness of using factor analysis with such a small sample (N=36) was also investigated.

Cluster analysis was effected using a simple technique called McQuitty's (1957) elementary linkage analysis (see Appendix 6.6). Working on the correlation matrix, each variable was assigned to the subset with which it had the highest single correlation. This produced subsets or clusters of variables with relatively high correlations among themselves and relatively low correlations with variables in the other clusters. There were a total of six groups or clusters (see Table 6.13). These six groups can be assumed to be sharing something in common - namely a common dimension or factor. According to Table 6.13, the first group seems to be concerned with 'composition', the second with 'attainment', the third with 'potential', the fourth with 'attitudes', the fifth with 'grading' and the sixth with 'appearance'.

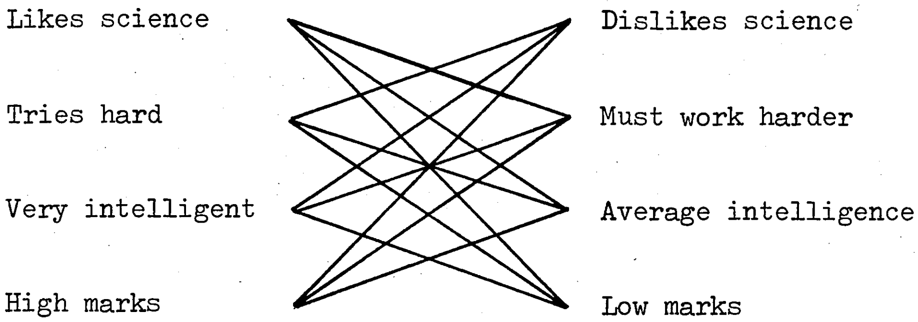
The cluster concerned with aspects of composition contains a number of variables and the dimension appears to have been over-sampled. Since the cluster is composed of variables that seemily refer to writing skills rather than to the content of the work samples, it was decided to drop some of the variables. Organisation of ideas, conciseness and completeness were chosen. Comparison of the mean ratings made of conciseness and completeness indicated that these two variables were ineffective at differentiating between the good and poor work samples. The variable conciseness had not been selected on an empirical basis during preliminary work (see section 6.9.1.4), and therefore it was of doubtful meaningfulness and relevance to the teachers. In addition, it correlated very poorly with the marks that the work samples were judged to merit. Completeness was also rejected on the basis of its low correlation with marks merited. Furthermore, the preliminary work had indicated that it was not one of the most popular factors used by science teachers to assess written work (see section 6.9.1.4). Organisation of ideas was largely rejected because of its high correlation with clarity of explanation. There was obviously considerable overlap between the meanings of the two variables. In addition, factor analysis showed organisation of ideas loading on a couple of factors. Thus the meaning of this variable was not simple. It was measuring more than one theoretical dimension.

6.9.1.7 Reliability

Reliability studies were conducted with sample ST(BIAS) (N=21). The final selection of rating variables yielded alpha coefficients of 0.77, 0.84, 0.86 when applied to the average, poor and good work samples respectively. A slightly higher value for coefficient alpha (0.87) was obtained for the good sample pair when CSE suitability was omitted from the calculation. This improved internal consistency resulting from the omission of CSE suitability provides an indication that the variable was inappropriate when applied to the good pupil (see also section 9.4.1.1).

Test-retest reliability was also calculated for a sub-sample of ST(BIAS) (N=14). The average correlation between the 15 ratings awarded to the good sample pair on the first and second administrations was 0.61.

Figure 6.1 The different attributes, and their combinations, that appeared on the cards



14/15 correlations for the individual variables were significant at the 5% level. One-tailed tests were used since positive correlations were predicted. The equivalent values obtained for the average and poor sample pairs were slightly lower.

6.9.2 Cards

The expectations that teachers form for male and female pupils when they are supplied with very limited information about those pupils was also investigated. Science teachers were presented with a pack of 30 cards, each of which described a different pupil, and they were asked to sort them according to their potential for GCE O level examinations in the science subjects.

Each card carried three facts about a pupil who was unnamed. One fact was the pupil's sex. This was conveyed by the noun 'he' or 'she'. The other two facts referred to work related attributes of the pupil. They were chosen from four different attributes: ability, attainment, effort and attitude to science. Combining positive and negative statements about these variables together in pairs produced 12 cards for each sex (see Figure 6.1). The 12 cards referring to girls exactly mirrored the 12 cards referring to boys. A further six cards, which did not form matching pairs, were included to introduce greater variety into the pack. These six cards linked pupil interest with a selection of the other variables.

The sorting exercise was introduced to the teachers in a very apologetic manner in order to secure their cooperation in what was a very artificial task. Furthermore, the undoubtedly tentative nature of their judgements was acknowledged (see Appendix 6.14). The teachers were asked to imagine that shortly after taking up a new teaching post, they were requested by their head of science to sort a class of third year pupils into potential O level candidates and those unsuited for the examination. The only information available to them was that contained in their predecessor's record book. The pack of cards represented that book.

6.9.2.1 First pilot

Sample A&A* were simply asked to sort the cards into two piles - one pile for potential GCE candidates and another for pupils not suited for this examination. Inspection of the results indicated that there was no evidence that the teachers were sex biased in their allocations. This was in spite of the fact that a number of the respondents sorted their cards in my presence, which prevented them from comparing any of cards before allocating them to the two piles.

6.9.2.2 Second pilot

In the second pilot the respondents (sample I) were asked to make finer discriminations between the pupils' potential for O level science. The teachers were first instructed to divide the cards into two piles - one for potential O level candidates and another for pupils not suited for the examination. Then they were asked to indicate how confident they were of each decision by writing:

'1' on the cards of those pupils they were very confident had been placed in the correct pile,

'2' on those cards that they were fairly confident had been placed in the correct pile,

'3' on the cards of those pupils whose allocation was somewhat uncertain.

This second sort arranged the cards along a 6-point ordinal scale.

The results of the second pilot are reported in Appendix 6.15. They show that pupil sex did have some effect upon the allocation of the cards, but that the effect was slight. This could have been due to teachers cheating in the exercise, i.e. first pairing the cards and then sorting the pairs. Close inspection of the results suggested that, in all probability, very few of the respondents had compared the cards before sorting them, but the possibility did exist. Because of this potential source of unreliability, the Cards exercise was not continued after the second pilot.

6.10 SCALE DEVELOPMENT DETAILS (E) FACT GATHERING

6.10.1 Personal Details

A series of questions were devised to gather factual information from respondents about themselves, their educational background, qualifications, past teaching experience and current teaching conditions. This information was primarily required to classify respondents, e.g. on the basis of their sex, principal teaching subject, teaching experience, etc. Some of the information, e.g. social class background, whether their mother worked, was needed to enable certain of the hypotheses to be tested. In addition, the opportunity was taken to gather descriptive data about science teachers, e.g. details regarding their qualifications, current teaching experiences.

As a result of the first pilot, given to sample A&A*, minor modifications were made to a few of the questions to clarify their meaning. A couple of the questions were expanded and one was simplified. After the second pilot (sample I), some questions were again simplified and rephrased, and a few, mainly those designed merely to collect descriptive data about science teachers, were discarded.

6.10.2 School Details

A number of details were requested from schools to enable correct classification of teachers according to the type of school in which they taught. Two different questionnaires were used. The shortest and simplest questionnaire collected details from the primary and middle schools involved in the COSS study. The other questionnaire, reproduced in Appendix 6.17, was designed for secondary schools and was sent to all schools that received BIAS and STOSS questionnaires. The questions requested general information about the nature of the school, e.g. the type of school, the size of the school, its location. The questionnaire passed through two pilot stages during its development, but only very minor changes were made as a result of the pilot work.

CHAPTER 7

RESULTS

Page

7.0 Contents

CHAPTER 7

TABLES

FIGURES

7.1 SEX TYPING OF SCIENCE

7.1.1 School Subject Characteristics

Replies were received from primary school teachers (sample O), middle school teachers (sample N) and secondary school teachers (sample M). The teachers indicated their responses on 7-point semantic differential rating scales. These ratings were subsequently converted to numerical data by assigning the value 1 to the rating position nearest that adjective of each adjective pair which pilot work had indicated was most closely associated with masculine/science qualities. This scoring procedure resulted in the rating position nearest the adjective that was most closely associated with feminine/arts qualities being assigned the value of 7. Thus low numbers denote stereotyped masculine characteristics, 4 signifies neutrality and higher numbers denote stereotyped feminine characteristics.

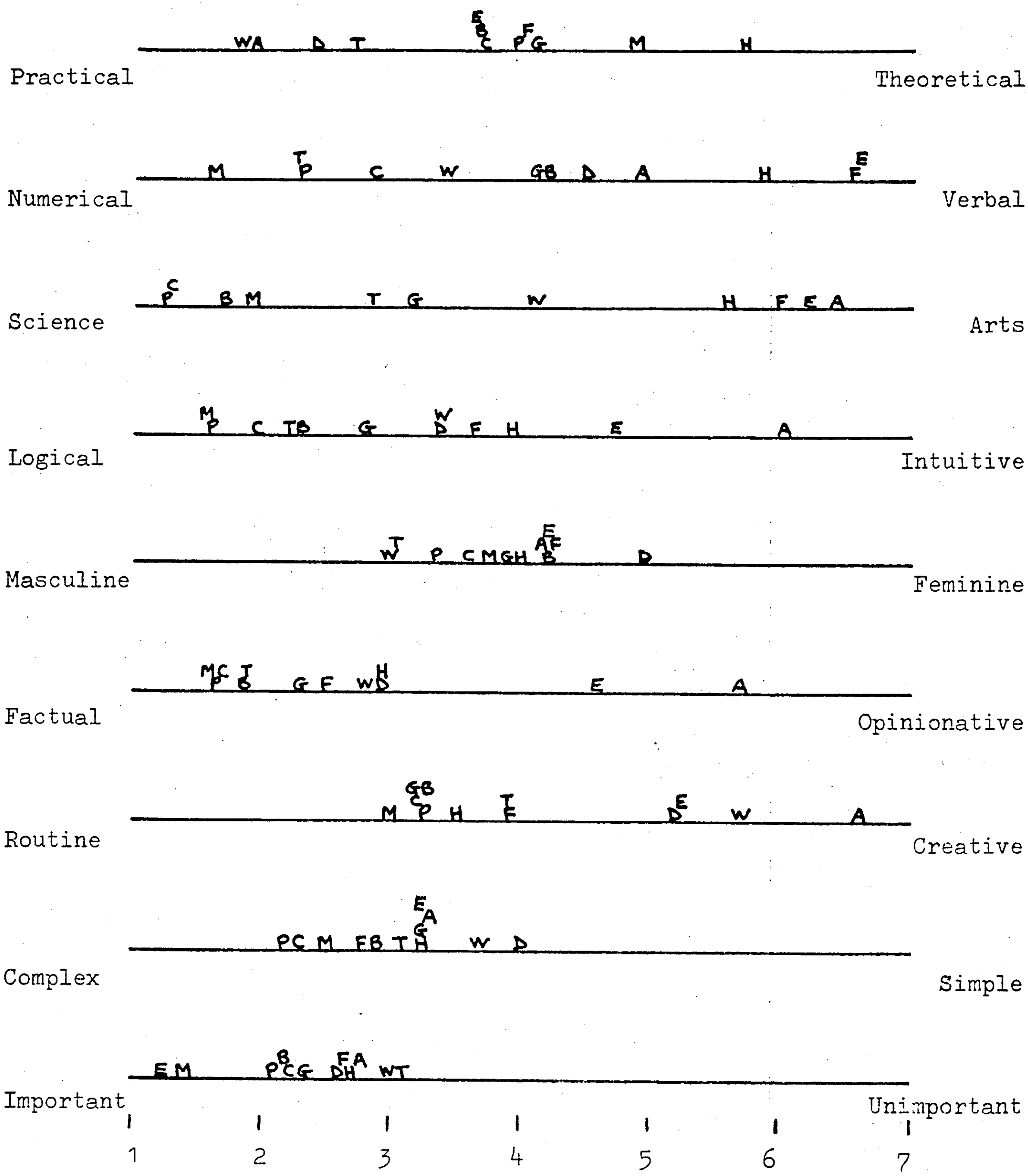
7.1.1.1 Science subjects compared with other subjects

The mean ratings received by each school subject on each rating scale from the whole sample (i.e. primary, middle and secondary teachers of all subjects) are recorded in Table 7.1. To facilitate comparisons, Figure 7.1 presents the same data in graphic form. It can then be seen that the subjects tend to fall into three groups:- science subjects (physics, chemistry, biology, maths), practical subjects (woodwork, home economics, technical drawing), and arts subjects (geography, history, French, English, art). These three groups of subjects appear in the same order on the numerical-verbal, science-arts, logical-intuitive and factual-opinionative scales. The science subjects have the lowest ratings, the practical subjects have intermediate ratings and the arts subjects have the highest ratings. The science subjects also have the lowest ratings on the routine-creative, complex-simple and important-unimportant scales. They received rather neutral ratings on the practical-theoretical scale. On the masculine-feminine scale, physics

Table 7.1 Mean ratings awarded to each subject on each dimension

	Physics	Chemistry	Biology	Maths	Home Economics	History	French	Woodwork	Geography	Art	English	Technical Drawing
Practical-theoretical	3.94	3.67	3.66	4.85	2.42	5.67	3.98	1.81	4.12	1.93	3.63	2.70
Numerical-verbal	2.27	2.85	4.18	1.61	4.48	5.84	6.54	3.40	4.09	4.90	6.62	2.25
Science-arts	1.25	1.28	1.69	1.88	3.37	5.57	6.01	4.11	3.16	6.43	6.23	2.86
Logical-intuitive	1.61	1.96	2.31	1.55	3.64	3.93	3.64	3.38	2.78	6.03	4.73	2.21
Masculine-feminine	3.36	3.59	4.20	3.76	4.96	3.99	4.24	2.97	3.89	4.16	4.18	3.03
Factual-opinionative	1.66	1.71	1.86	1.57	2.91	2.93	2.50	2.80	2.28	5.68	4.61	1.87
Routine-creative	3.24	3.21	3.27	2.97	5.17	3.53	3.93	5.71	3.18	6.61	5.25	3.91
Complex-simple	2.17	2.31	2.90	2.51	4.03	3.26	2.77	3.72	3.25	3.28	3.24	3.08
Important-unimportant	2.09	2.26	2.18	1.40	2.60	2.71	2.66	3.00	2.36	2.77	1.25	3.09

Figure 7.1 Mean ratings awarded to each subject on each dimension



Key

P	Physics	D	Home Economics	H	History
C	Chemistry	W	Woodwork	F	French
B	Biology	T	Technical drawing	G	Geography
M	Maths			A	Art
				E	English

Table 7.2 Science and non-science teachers' mean ratings of the science subjects

(A) Physics

Characteristic	Science teachers (N=53)	Non-science teachers (N=230)	t	p
Practical-theoretical	3.72	4.21	2.20	0.05
Numerical-verbal	2.00	2.06	0.41	ns
Science-arts	1.19	1.13	-0.82	ns
Logical-intuitive	1.53	1.57	0.30	ns
Masculine-feminine	3.32	3.29	-0.20	ns
Factual-opinionative	1.68	1.51	-1.45	ns
Routine-creative	3.85	3.15	-2.82	0.01
Complex-simple	2.06	2.16	0.67	ns
Important-unimportant	1.70	2.07	2.05	0.05

(B) Chemistry

Characteristic	Science teachers (N=53)	Non-science teachers (N=230)	t	p
Practical-theoretical	3.49	3.97	2.40	0.05
Numerical-verbal	2.94	2.69	-1.40	ns
Science-arts	1.26	1.17	-0.84	ns
Logical-intuitive	1.75	1,95	1.46	ns
Masculine-feminine	3.64	3.54	-0.82	ns
Factual-opinionative	1.62	1.57	-0.37	ns
Routine-creative	3.66	3.17	-2.07	0.05
Complex-simple	2.21	2.22	0.05	ns
Important-unimportant	1.96	2.34	2.49	0.05

(C) Biology

Characteristic	Science teachers (N=53)	Non-science teachers (N=230)	t	p
Practical-theoretical	3.58	3.84	1.31	ns
Numerical-verbal	4.23	4.16	-0.31	ns
Science-arts	1.62	1.56	-0.51	ns
Logical-intuitive	2.42	2.27	-0.83	ns
Masculine-feminine	4.32	4.25	-0.83	ns
Factual-opinionative	1.98	1.78	-1.31	ns
Routine-creative	3.83	3.11	-3.33	0.001
Complex-simple	2.83	2.94	0.59	ns
Important-unimportant	1.87	2.21	1.93	(0.10)

and chemistry were judged to be less masculine than woodwork and technical drawing, but more masculine than the arts subjects. Biology was judged to be very slightly feminine.

7.1.1.2 Effect of teaching subject upon perceptions of the science subjects

Science and non-science teachers' perceptions of a number of the characteristics possessed by the three main science subjects are compared in Table 7.2. The figures refer to the sample of secondary school teachers. By virtue of their teaching at the secondary level, these teachers were all equipped to express views about the science subjects as they are taught in secondary schools up to CSE/O level standard.

The science and non-science teachers generally viewed the science subjects very similarly, but there were some statistically significant differences (two-tailed t test). These differences are worthy of note since they tended to appear consistently across all three science subjects. The science teachers judged physics and chemistry to be more practical-based and less theory-based than did the non-science teachers. Furthermore, all three science subjects were considered to be more creative and less routine by the science teachers. Lastly, and not altogether surprisingly, the science teachers judged all three science subjects to be more important subjects than did the non-science teachers.

The views of the science subjects held by physics, chemistry and biology teachers were compared by one-way analysis of variance. Only one characteristic produced a significant result at the 5% level. Since this is less than could be expected to occur by chance, it can be concluded that physics, chemistry and biology teachers view each others subjects very similarly.

7.1.1.3 Effect of other teacher variables upon perceptions of the science subjects

Besides grouping teachers' replies according to the subject taught, two other teacher variables were also used to group the replies for comparative purposes. These variables were a teacher's sex, and the educational level at which a teacher taught.

Both male and female secondary science teachers viewed the science subjects very similarly. Of the 27 comparisons made between the mean ratings from male and female teachers (3 subjects x 9 characteristics), only three were statistically significant at the 5% level (two-tailed t test). This small proportion is hardly larger than would be expected to occur by chance. Therefore, it is concluded that male and female science teachers do not view school science subjects differently. This finding is further supported by statistical analysis (t test) of the replies to the same semantic differential scales received from sample TSCH. (The appropriateness of using evidence from another sample is discussed in Appendix 7.4). None of the 27 comparisons made between the mean ratings of male and female teachers were significantly different at the 5% level.

Male and female non-science secondary teachers' perceptions of the science subjects did differ slightly. The two most interesting differences and the ones most pertinent to this study were that male teachers regarded chemistry to be significantly more masculine and biology to be significantly more feminine than did female teachers. Full details of the mean ratings given by male and female non-science teachers are reported in Appendix 7.1, together with all the significant differences.

Differences between teachers' perceptions of the science subjects arising from the educational level at which they taught were investigated by one-way analysis of variance. Full results are recorded in Appendix 7.1. Generally, the secondary teachers gave the most extreme ratings and the primary teachers gave the least extreme ratings. Hence the secondary teachers sex typed the science subjects most. Physics was rated significantly more masculine and biology significantly more feminine by the secondary teachers than by the primary teachers.

7.1.1.4 Summary

1. Physics, chemistry, biology and maths formed a group of subjects that were generally viewed by teachers as being more numerical, scientific, logical, factual, routine, complex and important than a range of other secondary school subjects. They were judged to occupy an intermediate position on the practical-theoretical continuum. With the exception of biology, they received ratings on the masculine-feminine scale that were less masculine than woodwork and technical drawing, but more masculine than the arts subjects. Biology was judged to be very slightly feminine.

2. Secondary science teachers generally viewed the three science subjects very similarly to non-science teachers, except that science teachers judged physics and chemistry to be more practical subjects, and all three science subjects to be more creative and important than did non-science teachers.

3. Amongst other differences, male non-science secondary teachers regarded chemistry to be more masculine and biology to be more feminine than did female non-science teachers. Male and female science teachers did not view school science subjects differently.

4. One-way analysis of variance revealed that secondary teachers gave the most extreme ratings and primary teachers the least extreme ratings. Thus the secondary teachers sex typed the science subjects most by rating physics more masculine and biology more feminine.

7.1.2 Masculinity Index

Scores for the Masculinity Index were obtained by summing together the ratings on the four separate semantic differential scales for each science subject. Before summing, the polarity of two of the scales was reversed so that the masculine pole always received a low rating and the feminine pole a high rating. Since the semantic differential scale consisted of 7 rating positions, the lowest possible score over the four

Table 7.3 Masculinity Index values for the three science subjects (N=159)

Subject	Mean	S.D.	Minimum	Maximum
Physics	11.06	3.24	4.00	19.00
Chemistry	12.79	3.10	4.00	24.00
Biology	18.39	3.39	10.00	28.00

Table 7.4 Details of the Masculinity Index ratings given to the three science subjects (N=159)

(A) Physics and Chemistry

	Physics			Chemistry		
Adjective pair	Mean	S.D.	%LE3	Mean	S.D.	%LE3
Hard-soft	2.26	0.88	91.2	2.63	1.06	82.3
Tough-tender	2.67	1.03	75.5	3.09	1.04	62.0
Cold-warm	3.02	1.32	57.2	3.69	1.18	39.9
Remote-intimate	3.12	1.29	55.3	3.37	1.14	46.8

(B) Biology

Adjective pair	Mean	S.D.	%GE5
Hard-soft	4.09	1.38	36.9
Tough-tender	4.43	1.23	45.0
Cold-warm	4.99	1.17	65.0
Remote-intimate	4.88	1.11	58.8

%LE3 - Percentage of respondents giving a rating less than or equal to 3

%GE5 - Percentage of respondents giving a rating greater than or equal to 5.

scales was 4, and the highest possible score was 28.

Comparing the mean Masculinity Index scores received by the three science subjects (see Table 7.3), it can be seen that physics received a slightly lower score than chemistry, which means that it was considered to be the more masculine. However, it is clear that both physics and chemistry were regarded a masculine subjects. Several respondents even gave these two subjects the most masculine score possible. In contrast, biology received a slightly feminine score, but it was not so far displaced from a neutral score (16) as were the scores received by physics and chemistry.

Table 7.4 presents the mean ratings received by the three science subjects on the individual semantic differential scales. Also included are the percentage of respondents who gave the physical science subjects a masculine rating, i.e. a rating of 3 or less, on each scale and the percentage of respondents who gave biology a feminine rating, i.e. a rating of 5 or more, on each scale. The figures show that the masculine subjects, physics and chemistry, were viewed by most respondents as being 'hard'. In addition, over half of the respondents also associated physics with the masculine pole of the remaining three adjective pairs. The teachers' views about chemistry were not quite so polarised. Regarding biology, the teachers were in agreement that masculine adjectives were inappropriate. Biology only received a masculine rating for 'cold' from 8.1% of the respondents, and for 'remote' from 6.9%. In contrast, 65% of the respondents judged biology to be 'warm', a feminine rating.

7.1.2.1 Summary

1. The mean Masculinity Index scores given to physics and to chemistry indicated that the teachers judged these subjects to be masculine subjects. Physics was viewed as the more masculine of the two subjects. Biology received a slightly feminine score.

Table 7.5 Significance of difference between characteristics possessed by physics and biology (N=162)

Characteristic	Physics	Biology	t	2-tailed probability
Logical	1.33	2.09	-12.54	0.001
Objective	1.50	1.96	- 8.56	0.001
Relevant for careers	1.58	2.10	- 8.14	0.001
Relevant for family life	2.54	1.86	10.49	0.001
Mathematical	1.38	3.00	-28.95	0.001
Wordy	2.64	1.72	17.03	0.001
Concerned with people	2.94	1.85	16.41	0.001
Concerned with objects	1.48	2.69	-16.70	0.001
Concerned with social issues	3.08	2.20	12.56	0.001
Unfamiliar	2.36	2.96	- 9.02	0.001
Technical	1.54	2.61	-17.08	0.001
Mechanical	1.69	3.06	-20.68	0.001
Masculine	2.35	3.17	-12.43	0.001
Abstract	2.25	2.92	- 9.23	0.001
Impersonal	2.04	3.02	-12.27	0.001

2(a) More than 50% of the respondents rated physics as hard, tough, cold and remote. Thus physics received mean ratings towards the masculine pole on all four semantic differential scales comprising the Masculinity Index.

(b) Chemistry was rated hard and tough by more than 50% of the respondents.

(c) Biology was rated to be warm and intimate, both feminine ratings, by more than 50% of the respondents.

7.1.3 Characteristics of Science

The replies, which had been collected using a 4-point verbal rating scale, were converted to numerical data by assigning the value 1 to 'very' and so on up to 4 for 'not at all'. Thus high numbers denote that a characteristic is not linked with the subject, whilst low numbers imply that the adjective is very appropriate for the subject.

7.1.3.1 Comparisons between physics and biology

A simple paired t test was carried out on each item to compare the mean ratings assigned to physics and biology. The results appear in Table 7.5. It can clearly be seen that science teachers believe that physics possesses many characteristics that are very different to those possessed by biology. Physics was judged to possess significantly more of each of the listed characteristics, with the exception of wordiness, concern over social issues, concern over people, and relevancy for family life. These four characteristics were judged to be significantly better descriptors of biology. Thus, the teachers linked biology with features that are commonly associated with females, and physics was associated with many stereotypically masculine characteristics.

7.1.3.2 Characteristics associated with masculinity

To find out which characteristics of science subjects are perceived to be related to the gender connotations of the subject, Pearson product-

Table 7.6 Intercorrelations of science subject characteristics (Physics above the diagonal and biology below)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	-	51*	-02	03	22*	05	14	08	11	00	10	09	-08	00	06
2	41*	-	18*	-01	14	04	11	25*	14	-02	26*	11	01	14	03
3	18*	20*	-	18*	-08	04	21*	08	34*	-02	09	02	15	00	-06
4	14	12	33*	-	-06	02	41*	-06	40*	-16*	00	-05	-03	-12	-40*
5	38*	20*	20*	19*	-	01	-08	25*	07	26*	32*	34*	06	05	18*
6	-11	-11	-12	-09	-16*	-	10	04	06	00	08	03	01	03	-04
7	15	00	10	53*	21*	-03	-	-13	45*	-10	-07	03	-12	-06	-36*
8	09	-07	17*	-02	22*	-01	03	-	09	-02	25*	33*	15	15	15
9	11	02	22*	44*	14	06	39*	15	-	-09	04	07	-02	-02	-25*
10	-03	01	-11	-26*	02	11	-15	12	-06	-	27*	19*	15	07	35*
11	16*	16*	09	08	22*	10	07	18*	14	15	-	54*	20*	03	14
12	30*	04	15	-01	39*	01	07	32*	19*	07	33*	-	24*	-03	11
13	-05	02	07	-01	04	16*	-07	09	16*	12	15	27*	-	18*	19*
14	03	06	-12	-22*	02	04	-22*	21*	-05	31*	18*	20*	10	-	35*
15	-09	05	-07	-49*	00	08	-53*	10	-36*	26*	05	09	14	27*	-

Key

1 Logical
2 Objective
3 Relevant for careers
4 Relevant for family life
5 Mathematical
6 Wordy
7 Concerned with people
8 Concerned with objects
9 Concerned with social issues
10 Unfamiliar
11 Technical
12 Mechanical
13 Masculine
14 Abstract
15 Impersonal

Note Decimal points omitted
* significant at 5% level (2-tailed test)

moment correlation coefficients were calculated between the variable 'masculine' and each other variable for both physics and for biology. The correlations appear in Table 7.6, together with the intercorrelations between all the other variables. The masculine aura of physics is linked most strongly with the perception that physics is technical and mechanical (p 0.01). Likewise, the degree to which biology is viewed as being mechanical is also related to the degree to which the subject is judged to be masculine.

The dimensions underlying the interrelationships between the different variables were investigated using elementary linkage analysis (see Appendix 6.6). When applied to physics, the 15 variables formed three clusters, whilst four clusters emerged for biology. The composition of these clusters is shown in Table 7.7, together with an indication of whether the characteristic was associated with the subject or not. Characteristics which received a mean rating of less than 2.5, i.e. the subject was thought to be 'very' or 'fairly' logical, objective, etc., are deemed to be associated with the subject; whilst characteristics which received a mean rating of greater than 2.5, i.e. the subject was thought to be 'not very' or 'not at all' logical, objective, etc., are considered to be not associated with the subject.

The cluster which refers to the masculine image of physics, besides containing the adjective 'masculine', also includes the following characteristics: technical, mechanical, mathematical, and concerned with objects. All of the intercorrelations are positive. The cluster for biology referring to gender image contains the same characteristics, with the addition of 'wordy'.

7.1.3.3 Summary

1. Science teachers believe that physics and biology, as they are taught in secondary schools up to CSE/O level standard, differ on a number of characteristics. The two subjects were given significantly different

Table 7.7 Some characteristics of two science subjects, clustered by elementary linkage analysis

Physics (N=164)	Biology (N=162)
1. Technical (+)	1. Concerned with social issues (+)
Mechanical (+)	Relevant for careers (+)
Mathematical (+)	Relevant for family life (+)
Concerned with objects (+)	Concerned with people (+)
Masculine (+)	Impersonal (-)
2. Logical (+)	2. Logical (+)
Objective (+)	Objective (+)
3. Unfamiliar (+)	3. Technical (-)
Abstract (+)	Mechanical (-)
Impersonal (+)	Mathematical (-)
Relevant for family life (-)	Concerned with objects (-)
Wordy (-)	Masculine (-)
Concerned with people (-)	Wordy (+)
Concerned with social issues (-)	4. Unfamiliar (-)
Relevant for careers (+)	Abstract (-)

(+) Subject associated with characteristic

(-) Subject not associated with characteristic

Table 7.8 Science teachers' ranking of the most important factors that give the physical science subjects a masculine image (N=35)

Public Opinion	Personal Opinion
1. Number of male scientists	1. Adverts
2. Mechanical aspects	2. Films
3. Tradition	3= Comics
4. Stereotyping	3= Stereotyping
5. Social pressure	3= Social pressure

ratings for all the 15 characteristics investigated. Physics was judged to possess more of those characteristics that are commonly associated with males, e.g. logical, mathematical, technical, mechanical; whilst biology was linked with characteristics that are commonly associated with females, e.g. wordiness, concern over social issues, concern over people.

2. The belief that physics can be described as 'masculine' correlates significantly with the belief that physics is a technical and mechanical subject.

3. Elementary linkage analysis produced a clearly defined 'masculine' cluster that contained five characteristics: masculine, technical, mechanical, mathematical, and concern with objects. All of the inter-correlations between them were positive.

7.1.4 Opinions

7.1.4.1 The causes of science's masculine image

The replies received from samples F and H were pooled (the samples were very similar, both being secondary science teachers in mixed comprehensive schools) to produce a combined sample of 35 respondents. These science teachers' own opinions and their perceptions of the general public's opinion as to which factors give the physical science subjects a masculine image are recorded in Figures 7.2 and 7.3. The figures show the number of teachers who indicated that 'yes' an item does contribute to the masculine image of the physical science subjects or 'no' it does not. Teachers who responded by using the '?' category have been excluded. Their numbers were very small, and they did not contribute significantly to the overall pattern of replies.

Table 7.8 ranks the five items from each of public opinion and personal opinion that received the highest 'yes' scores, after the 'no' scores had been subtracted. The table shows that the teachers believe that the public associates the masculine image of the physical science subjects with the preponderance of male scientists, and the presence of

Figure 7.2 Perceived public opinion about the causes of the masculine image of physical science subjects

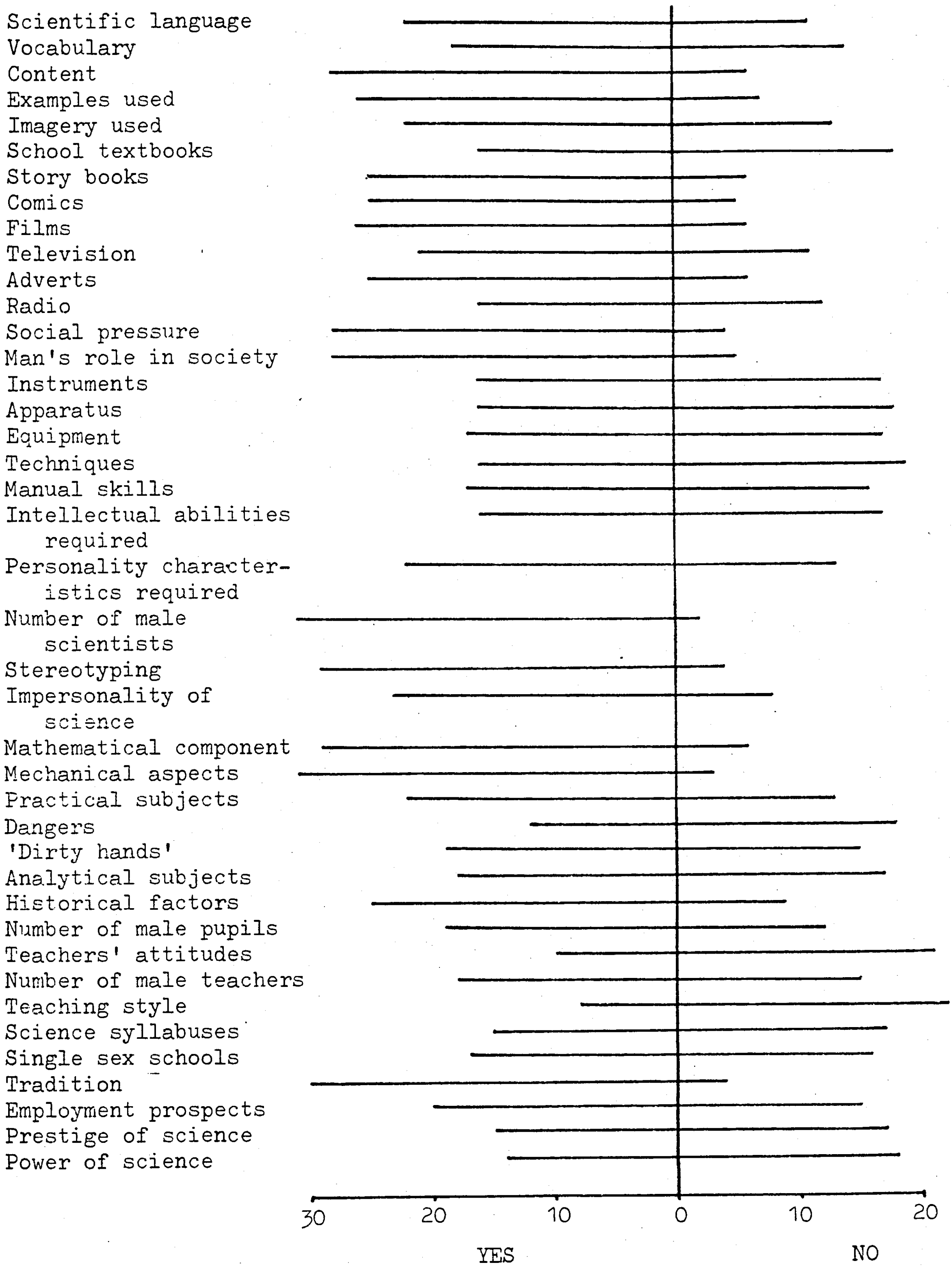


Figure 7.3 Science teachers' personal opinions about the causes of the masculine image of physical science subjects

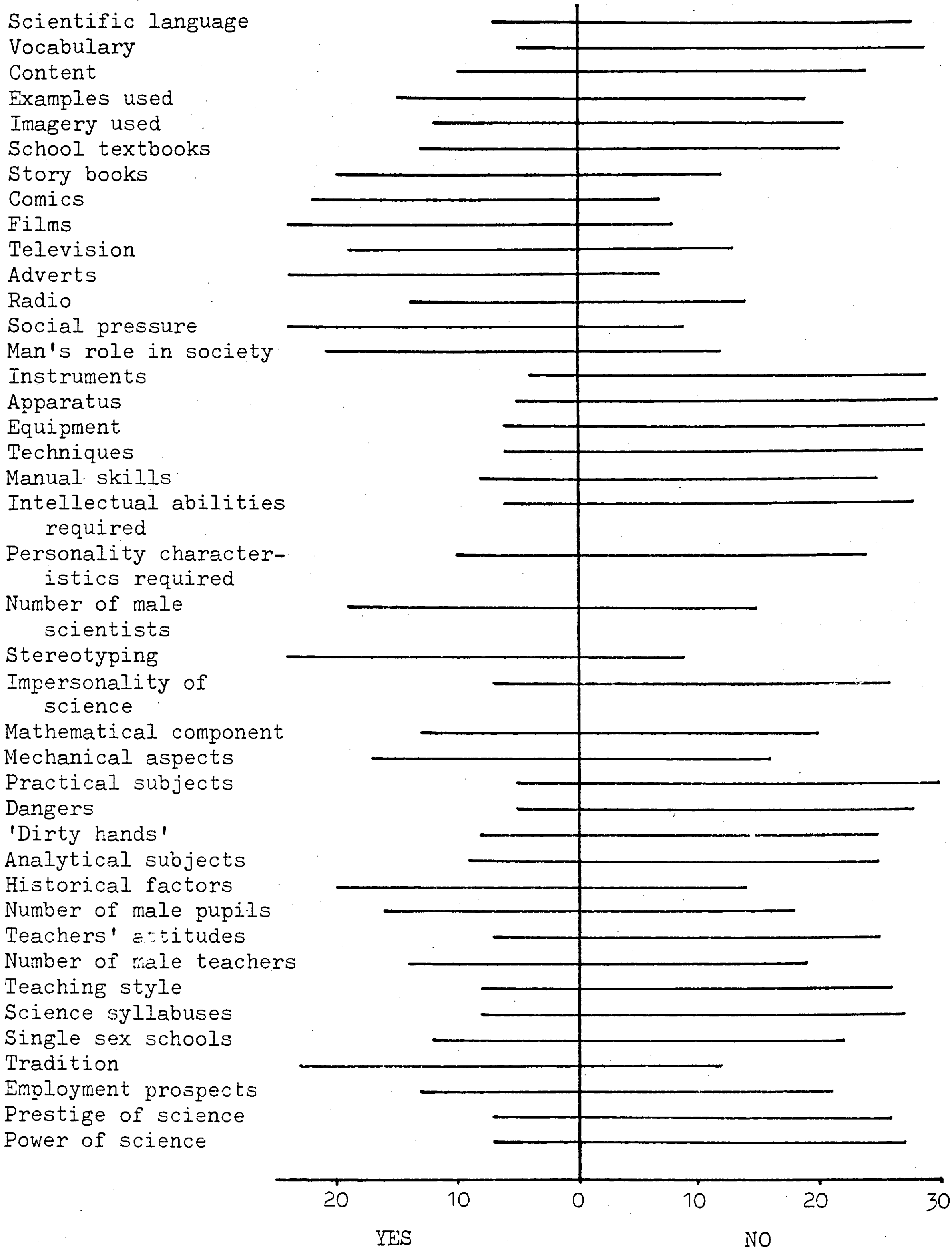


Table 7.9 Science teachers' ratings of the gender image of physical and biological science on a simple scale

(A) PUBLIC OPINION

	Physical Science			Biological Science		
Sample	Masculine	Neutral	Feminine	Masculine	Neutral	Feminine
C	13	1	-	-	8	6
F	12	2	-	-	7	7
H	18	4	-	-	6	16

(B) PERSONAL OPINION

	Physical Science			Biological Science		
Sample	Masculine	Neutral	Feminine	Masculine	Neutral	Feminine
C	2	13	-	-	12	3
F	6	8	-	-	11	3
H	2	20	-	-	21	1

mechanical elements in physical science. Both of these factors are closely entwined with physical science in its present form. The remaining three factors in Table 7.8 all refer to the vague pressures which emanate from the ttitude and expectations commonly found in our society.

In contrast, the teachers were of the opinion that various facets of the media are instrumental in promoting a masculine image for physical science. However, they did also indicate that social pressures are important in maintaining science's masculine image.

7.1.4.2 The gender of science

Replies from all three samples (C,F,H) to the direct question about the gender image of the science subjects are recorded in Table 7.9. The upper half of the table shows the opinions that science teachers believe are to be found within the general public. The lower half of the table contains the teachers' own perceptions regarding the gender of science subjects. The table clearly shows that the teachers believe that the general public does view the physical science subjects as being masculine subjects, whilst if any gender attribute is associated with the biological sciences, it is femininity. In contrast, the teachers maintained that they view both the physical and biological sciences as neutral subjects.

7.1.4.3 Summary

1. Science teachers think that the general public regard the science subjects, as they are taught in secondary schools up to CSE/O level standard, to be gender linked. The physical science subjects, e.g. physics and chemistry, are viewed as being masculine subjects, and the biological science subjects as being slightly feminine subjects.

2. In answer to the same direct question, science teachers maintained that they view both the physical and biological sciences as neutral subjects.

3. Science teachers believe that the public primarily associate the

Table 7.10 Science teachers' perceptions of scientists (N=164)

	Physicist	Biologist	t	p
Male	2.38	4.42	-18.01	0.001
Good at maths	1.63	3.95	-23.61	0.001
Logical	1.86	3.15	-12.11	0.001
Objective	2.11	2.82	- 7.05	0.001
Competitive	3.15	3.52	- 4.01	0.001
Unsociable	4.66	5.20	- 4.31	0.001
Unemotional	4.30	4.96	- 5.61	0.001
Not humanitarian	4.27	5.04	- 6.63	0.001

masculine image of the physical science subjects with the preponderance of male scientists, and the presence of mechanical aspects in physical science. Public opinion is also believed to regard tradition, stereotyping and social pressures as important factors contributing to science's masculine image.

4. The personal opinion of science teachers is that the media, e.g. adverts, films, comics, are primarily responsible for giving the physical science subjects a masculine image. They suggest that stereotyping and social pressures are important contributory factors helping to maintain this image.

7.1.5 Scientist Stereotypes

Replies were coded such that the lower the number, the greater the probability that the scientist possesses the characteristic. Conversely, the higher the number, the more improbable it is that the scientist possesses that characteristic.

Table 7.10 compares the mean ratings awarded to a physicist and a biologist. The statistical significance of each difference was ascertained using a paired t test. The t values are recorded in Table 7.10, together with the associated p values for a two-tailed test. It can be seen that all of the differences are highly significant, and that sex produced one of the largest differences between mean ratings.

Science teachers clearly differentiate between different types of scientists. The teachers indicated that a physicist was more likely than a biologist to possess all of the eight characteristics investigated. These characteristics included an ascribed attribute (sex, i.e. male), attributes that scientists are known to possess (e.g. objectivity), characteristics stereotypically associated with scientists (e.g. unsociability), and characteristics stereotypically associated with men (e.g. competitiveness).

7.1.6 Conclusions

1. Physics, chemistry, biology and maths form a closely related group of school subjects that can broadly be termed 'science' subjects. Teachers regard these subjects to be more numerical, scientific, logical, factual, routine, complex and important than a range of other secondary school subjects (Figure 7.1).

2(a) The use of direct and explicit measuring scales to determine science teachers' views about the gender connotations of the science subjects tends to produce neutral ratings. Similarly, the use of scales with few rating positions tends to produce neutral ratings. Measuring scales that are less transparent and/or have more rating positions tend to produce more differentiated responses.

(b) Discriminating scales indicate that science teachers regard physics and chemistry to be masculine subjects. This finding supports Hypothesis One. Science teachers perceive physical science to have a masculine image. In contrast, biology is viewed as a very slightly feminine subject (Table 7.3).

p 100

(c) The results from two scales (School Subject Characteristics and Masculinity Index) indicated that science teachers judge physics to be more masculine than chemistry. It has already been noted that biology is judged to be slightly feminine. Thus Hypothesis Two is confirmed. Science teachers rank the three common science subjects in order of masculinity - physics, chemistry, biology - with physics being the most masculine subject.

3. (Relates to Question Three, How do the views of secondary science teachers about the masculinity of science compare with those of teachers of other subjects and teachers from other educational levels?) Science teachers' views about the gender connotations of the three science subjects are not significantly different to the views of other secondary teachers (Table 7.2). However, secondary teachers (including science

teachers) view physics to be significantly more masculine than do primary teachers. They judge biology to be significantly more feminine than do primary teachers (Table A7.1/2, Appendix 7.1). Thus secondary teachers are more inclined than primary teachers to sex type the science subjects.

4. (Relates to Question Two, What subject characteristics are associated with a masculine image?) The masculine image of physics is linked in teachers' minds with a number of other characteristics that the subject is believed to possess. Elementary linkage analysis produced a clearly defined 'masculine' cluster that contained five characteristics: masculine, technical, mechanical, mathematical, and concern with objects (Table 7.7).

5(a) (Relates to Question One, How do science teachers account for the masculine image of science?) Science teachers express the view that the media, e.g. adverts, films, comics, are primarily responsible for giving the physical science subjects a masculine image. They believe that stereotyping and social pressures also play a role (Table 7.8).

(b) Science teachers believe that the public associates the masculine image of the physical science subjects with the preponderance of male scientists, and the presence of mechanical elements in physical science. Public opinion is also believed to regard tradition, stereotyping and social pressures as important factors contributing to science's masculine image (Table 7.8).

5(a) Science teachers believe that a physicist is quite probably male. They also judge that a physicist is good at maths, logical, objective and competitive. Besides being appropriate characteristics for a scientist, all of these characteristics are also stereotypically associated with men. Science teachers are uncertain about the likihood of a physicist being unsociable, unemotional and not humitarian (Table 7.10). These findings confirm Hypothesis Four, which refers to science teachers' perceptions of physical scientists. Science teachers mainly associate scientists with

Table 7.11 Percentage of teachers who believe that they can generally distinguish between the written work of boys and girls

	Men	Women	Total
Science teachers	79.7	56.7	71.9
Non-science teachers	79.5	68.4	73.3
Total	79.6	64.4	72.8

masculine qualities and rarely with feminine qualities.

(b) Science teachers believe that a biologist is equally likely to be male or female. Furthermore, they believe that a biologist is significantly less likely than a physicist to display the above mentioned characteristics (Table 7.10).

7.2 SEX STEREOTYPING

7.2.1 Written Work of Girls and Boys

The replies from all the samples who answered questions about the written work of girls and boys were pooled together giving 89 replies from science teachers and 101 replies from non-science teachers. An initial perusal of the features that were associated with the written work of girls and boys indicated that although the replies of science and non-science teachers were in many respects very similar, there were a few obvious diffrerences in their emphasis. Therefore it was decided to analyse the replies of science and non-science teachers separately. Since the main focus of this study is upon science teachers, greater attention is devoted to the replies of the science teachers. The replies of the non-science teachers provide a useful index of the beliefs of a wide range of teachers, against which the results from the science teachers can be compared to assess their typicality.

7.2.1.1 The ability of teachers to recognise the work of each sex

73% of all teachers gave an affirmative answer to the question 'Would you say that you can generally distinguish between the written work of boys and girls?" A breakdown of the replies from different groups of teachers appears in Table 7.11. Overall, the responses of men and women were significantly different ($\chi^2 = 5.51$, $p<0.05$), with a higher proportion of men answering in the affirmative. This sex difference largely arose because of differences between the replies of male and female science

Table 7.12 Percentage of teachers who can recognise differences between the written work of boys and girls

	Men	Women	Total
Science teachers	84.7	86.7	85.4
Non-science teachers	86.4	78.9	82.2
Total	85.4	81.6	83.3

teachers. A significantly higher proportion of male than female science teachers gave an affirmative answer ($\chi^2 = 4.13$, $p<0.05$), whilst the replies of male and female non-science teachers were not significantly different. Table 7.11 also shows that a lower percentage of female science teachers than female non-science teachers gave affirmative answers, but the difference is not statistically significant. A detailed breakdown of the science teachers replies by teaching subject and sex appears in Appendix 7.2. Although certain trends are discernible, notably that biology teachers were less likely to give affirmative answers than physical science teachers, the small numbers of teachers involved, especially female teachers, precludes statistical analysis.

40% of all the teachers who indicated that they could not distinguish between the written work of boys and girls, nevertheless went on to list differences between boys' and girls' written work. Thus a higher proportion of the teachers (83%) indicated that they recognise differences between boys' and girls' written work, than were prepared to claim that they could use those differences to distinguish between the written work of boys and girls (73%). Table 7.12 presents the percentage of different groups of teachers who listed differences between girls' and boys' written work. In contrast to the teachers' claims to be able to distinguish between the written work of boys and girls (see Table 7.11), there are no significant differences between the replies of men and women, or science and non-science teachers. A detailed breakdown of the science teachers' replies by teaching subject and sex appears in Appendix 7.2. There was greater agreement amongst the replies than there was over the teachers' claims to be able to distinguish between the written work of boys and girls.

7.2.1.2 Features associated with the work of each sex

The second question "Can you briefly indicate any features that you consider to be typical of the written work of girls and boys" produced

Table 7.13 Frequency (%) with which teachers refer to different aspects of written work when listing differences between boys' and girls' work

		Science teachers (N=76)	Non-science teachers (N=83)
Handwriting			
	M	22.0	39.5
	F	23.1	22.2
Appearance			
	M	98.0	76.3
	F	84.6	86.7
Approach			
	M	34.0	31.6
	F	30.8	31.1
Content - Language			
	M	6.0	7.9
	F	7.7	8.9
Content - Style			
	M	8.0	23.7
	F	15.4	17.8
Content - Appraisal			
	M	54.0	21.1
	F	38.5	17.8
Content - Quantity			
	M	26.0	18.4
	F	34.6	13.3

M - male teachers

F - female teachers

a range of replies. To facilitate the examination and interpretation of these diverse responses, they were organized and categorized.

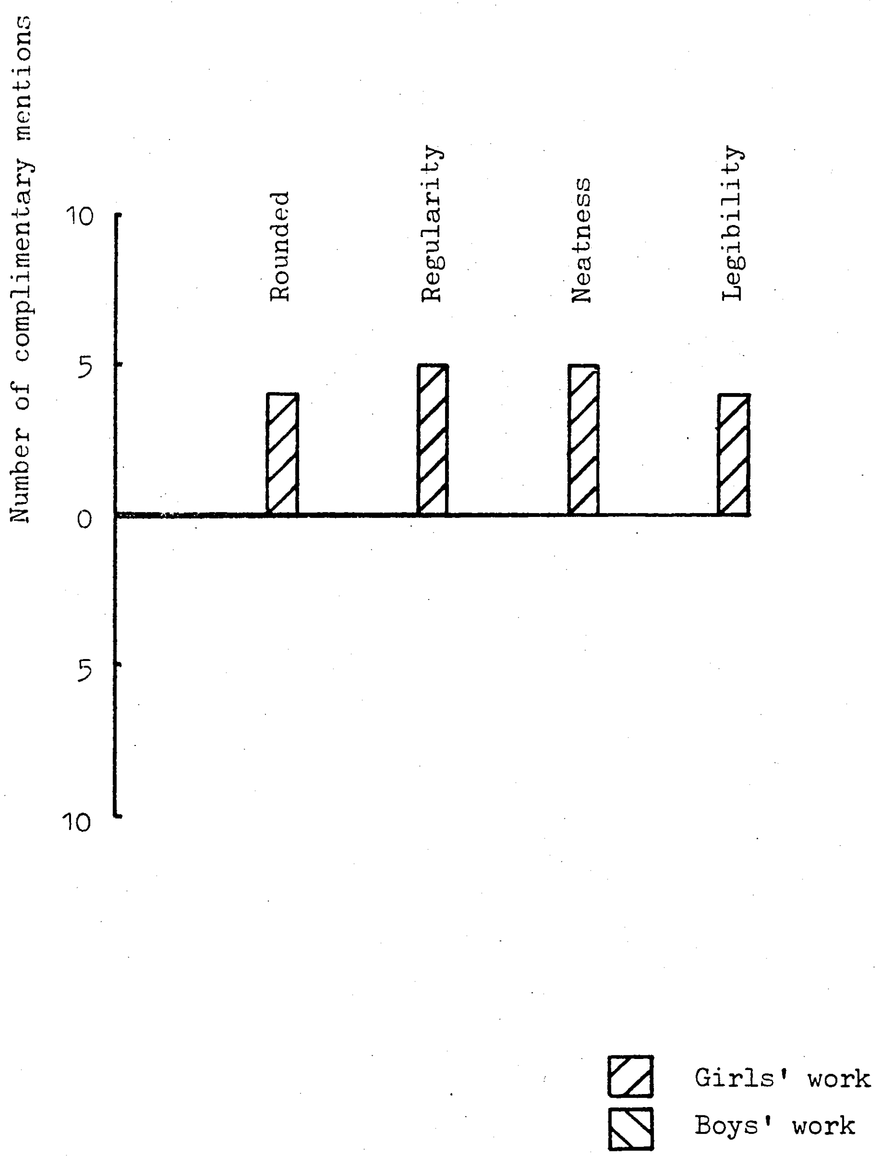
The replies were first listed according to the terms used by the teachers themselves. Then any similar terms were combined and an encompassing descriptor was supplied. For example, concise, short, brief, scant, laconic, succinct, economical, terse and minimum possible were all recorded under the descriptor 'brief'. Finally all the features referring to a particular aspect of written work were grouped together under one heading. For example, neatness, legibility, size and regularity of handwriting style were grouped with the different styles mentioned under the general heading of 'handwriting'. This exercise resulted in the formation of four major groups of features - handwriting, appearance, approach and content. The last category encompasses four aspects of content: the language in which the piece of work is written, the style of writing, the amount of writing and an appraisal of its worth.

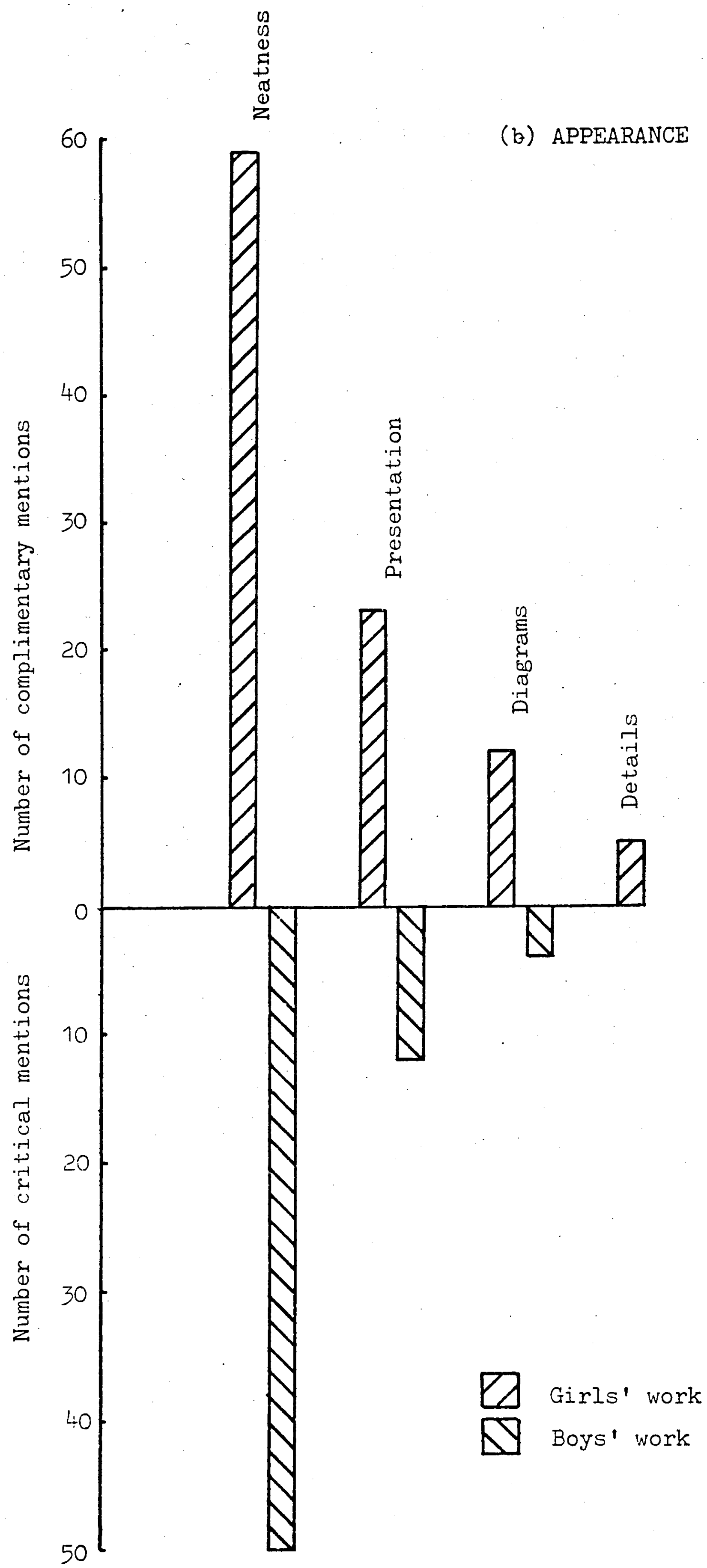
Figure 7.4 shows the range of features mentioned by the science teachers, and the manner in which they were grouped into the four main categories. It also indicates the number of science teachers who mentioned each feature and whether that mention was in a critical or complimentary context. Naturally the results for boys and girls are shown separately. Detailed results from the non-science teachers appear in Appendix 7.2, together with the results from the science teachers expressed in percentages for ease of comparison with the non-science teachers' results. Features which were mentioned by less than 5% of those teachers who listed differences have not been recorded.

Of those teachers who listed differences between the written work of boys and girls, the percentage mentioning features within each of the major categories is shown in Table 7.13. The range of features mentioned by science and non-science teachers was broadly similar, but there were several notable differences. These are discussed further in the

Figure 7.4 Features that are viewed as typical of girls' and boys' written work by science teachers

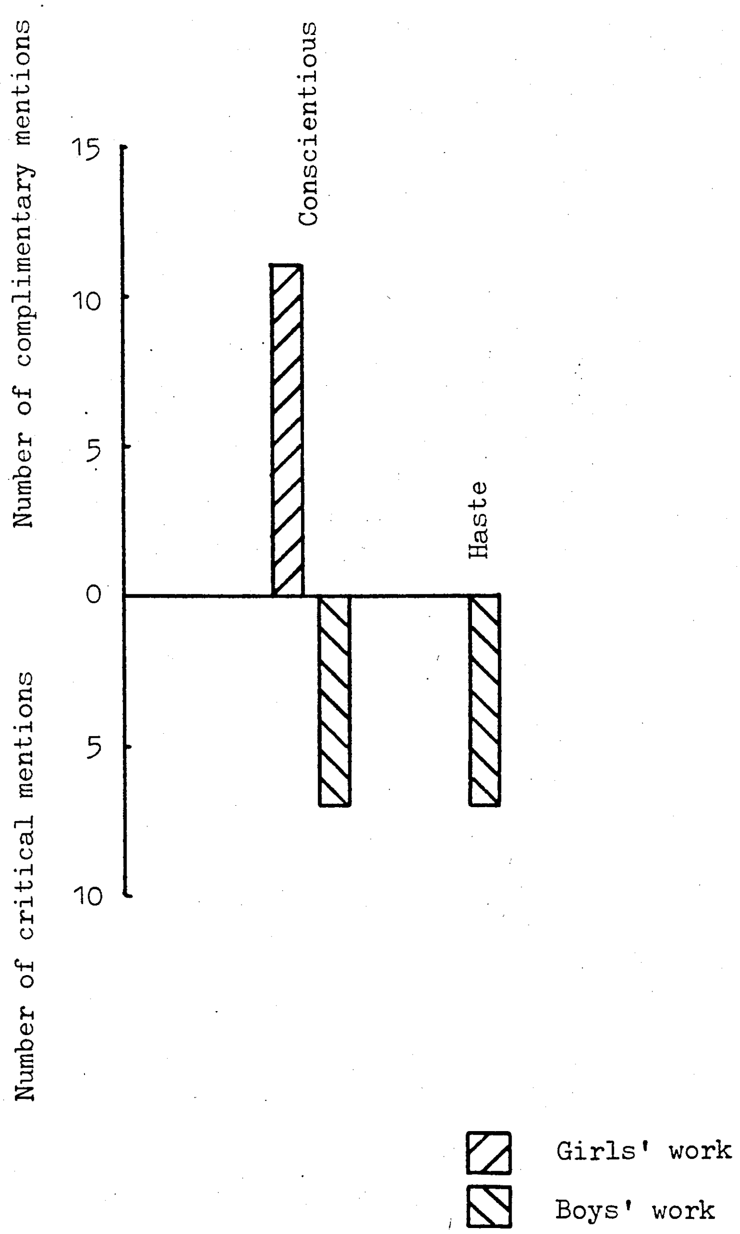
(a) HANDWRITING



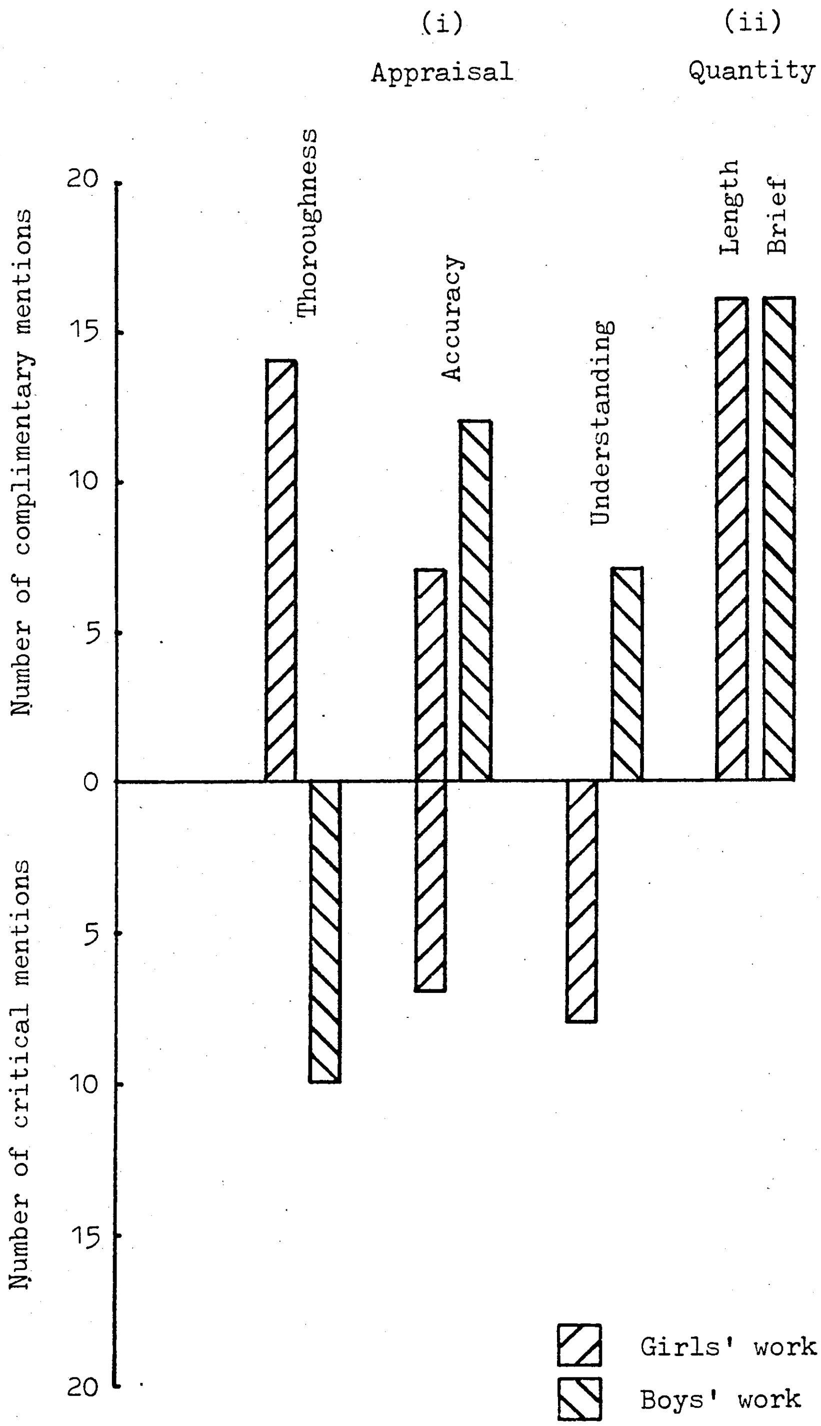


(b) APPEARANCE

(c) APPROACH


Number of complimentary mentions
Number of critical mentions
15
10
5
0
5
10
Conscientious
Haste
Girls' work
Boys' work

(d) CONTENT



(i)
Appraisal
(ii)
Quantity
Thoroughness
Accuracy
Understanding
Length
Brief
Number of complimentary mentions
Number of critical mentions
20
15
10
5
0
5
10
15
20
Girls' work
Boys' work

appropriate paragraphs below.

Over 90% of those science teachers who listed differences between the written work of boys and girls mentioned that the appearance of girls' work differs from that of boys'. A massive 98% of male science teachers mentioned differences in appearance. This response rate was marginally higher than that encountered among female science teachers ($\chi^2 = 3.05$, $p<0.10$), and significantly higher than that encountered among male non-science teachers ($\chi^2 = 8.04$, $p<0.01$). The feature most commonly mentioned by both science and non-science teachers was neatness. The teachers were unanimous in their judgement that the work of girls tends to be neat and tidy, whereas that of boys tends to be untidy. Good presentation was frequently associated with girls's work, but never with boys' work. A number of science teachers mentioned that girls produce diagrams of a much higher standard than do boys. The diagrams of girls are "drawn with a rule, neat, large and sometimes coloured in", whereas boys produce "smaller (more scruffy) diagrams". Girls also pay particular attention to details, e.g. "date, underlining titles, ruling off work".

The difference in appearance between girls' and boys' work was often attributed to their differing approaches to work. Girls were described as being conscientious, painstaking, careful, fussy and meticulous. But boys were seen as being careless. One teacher described boys as having "a can't care less attitude about presentation". Several teachers mentioned that boys' work conveys the impression that it has been produced in haste. A physics teacher wrote that boys' written work is "often rushed in order to get back onto practical work".

Aspects of handwriting were mentioned by over 20% of science teachers who listed differences between the written work of boys and girls. The writing of girls was judged to be neat, large and easily legible. The style was most often described as upright, rounded and regular. Particular features of boys' handwriting were mentioned by less than 5% of both science and non-science teachers.

The language used by boys was generally considered to be poorer than that used by girls. Specific points mentioned included vocabulary, punctuation, spelling and grammar. The style of writing adopted by boys and girls was mentioned by 21% of the non-science teachers, but by only 11% of the science teachers. No one feature of writing style was mentioned by 5% of the science teachers. In contrast, at least 5% of the non-science teachers judged the style of boys' writing to be more objective and imaginative than that of girls, whilst girls' writing was seen as being more expressive. Opinions were divided regarding the presence or absence of imaginative thought in girls' work.

Nearly 50% of all science teachers who listed differences between boys' and girls' written work commented upon the quality of the work produced by each sex, whilst less than 20% of non-science teachers made such comments. This difference between science and non-science teachers in their tendency to refer to aspects of work quality when asked to give features typical of boys' and girls' work is highly significant ($\chi^2 = 15.44$, $p<0.001$). In their appraisal, science teachers were both critical and complimentary about the work of each sex. Girls' work was judged to be much more thorough (detailed, comprehensive, complete) than that of boys. There was agreement that the work of boys is accurate, but there was disagreement over the accuracy of girls' work. One science teacher wrote about boys' written work, "Broad outlines sound, detail sometimes lacking or garbled (whole greater than the sum of the parts)". A chemistry teacher described boys' work as being "often accurate and to the point". Besides expressing greater confidence over the accuracy of boys' work, science teachers also credited it as showing more understanding than that of girls. According to a chemistry teacher, a boys' written work "often shows a greater depth of understanding". In contrast, girls are more likely to produce "regurgitated lesson notes or textbook notes written without understanding" (physics teacher).

There was complete agreement over the amount of work that boys and girls produce. All mentions in this sub-category referred to the length of girls' work and the brevity of boys' work. Boys' work was "brief, but usually covering the essential", whereas girls' work was "wordy, often with unnecessary detail". It is interesting to note that science teachers made significantly more mentions about the quantity of work produced by boys and girls than did non-science teachers ($\chi^2 = 4.08$, $p<0.05$).

Finally non-science teachers, but not science teachers, referred to the topics written about by boys and girls. A religious education teacher wrote that girls "tend to write about human and social issues rather than about mechanistic or philosophical matters". Other teachers indicated that girls write about romance, 'fairyland', happy endings, parents, magic. Boys write about more adventurous topics and "'masculine' matter". Topics mentioned included war, fights, heroes, science fiction, space, travel, sport and crime.

7.2.1.3 <u>Summary</u>

1. 73% of teachers of all subjects claimed that they can generally distinguish between the written work of boys and girls. Among science teachers, a significantly higher proprtion of male than female teachers claimed to be able to recognise the work of each sex.

2. 83% of teachers of all subjects were able to list features that they consider to be typical of the written work of each sex.

3. The features mentioned by science teachers who listed differences between the written work of boys and girls fell into the following categories.

(a) Over 90% of the teachers noted differences in the appearance of girls' and boys' written work. There was unanimous agreement that girls' work tends to be neat and well presented, with good diagrams. In contrast, boys' work was described as being untidy and poorly presented.

(b) Over 30% of the teachers perceived differences between the sexes in their approach to their work. Girls were seen as being conscientious and boys as being careless.

(c) Aspects of handwriting were mentioned by over 20% of the science teachers. They judged girls' writing to be rounded, neat and easily legible.

(d) Nearly 50% of the teachers commented upon the quality of the work produced by each sex. Boys were credited as displaying more understanding than girls, and the teachers were agreed that boys produce accurate work. There was disagreement over the accuracy of girls' work, but not over its thoroughness.

(e) About 30% of the teachers noted the amount of work produced by boys and girls. All mentions referred to the length of girls' work and the brevity of boys' work.

7.2.2 Preference for Subject Characteristics

Replies were received from primary school teachers (sample O), middle school teachers (sample N), secondary school teachers of various subjects (sample M), and two groups of science teachers - those comprising the main study (sample P), and a smaller group of teachers from eight schools where the pupils also filled questionnaires (sample TSCH). All of the teachers indicated their responses on 7-point semantic differential rating scales. These ratings were subsequently coded as numerical data (see section 7.1.1 for details). Low numbers denote stereotyped masculine/science characteristics, 4 signifies neutrality and higher numbers denote stereotyped feminine/arts characteristics.

The results presented below focus upon the replies received from sample P, the science teachers in the main study. However, reference is also made to the replies received from the other samples, as they allow the replies from the science teachers to be placed in a wider context. The validity of comparing results from several different sources is discussed in Appendix 7.4.

Table 7.14 Pupil preference for subject characteristics as seen by science teachers (N=159)

Characteristic	Mean Rating Boy	Mean Rating Girl	t	p
Practical-theoretical	2.80	3.45	5.83	0.001
Numerical-verbal	3.64	5.24	15.03	0.001
Science-arts	2.94	4.43	13.60	0.001
Logical-intuitive	3.14	4.34	10.75	0.001
Masculine-feminine	2.43	5.04	17.27	0.001
Factual-opinionative	3.09	3.84	5.66	0.001
Routine-creative	4.17	4.48	2.49	0.050
Complex-simple	4.25	4.64	4.28	0.001
Important-unimportant	2.79	2.94	2.13	0.050

7.2.2.1 The views of science teachers

Science teachers perceptions of the subject characteristics preferred by boys and girls are recorded in Table 7.14. The statistical significance of the difference between each pair of mean ratings was investigated by conducting paired t tests. The t values obtained are shown in the table, together with indications of their significance under two-tailed test conditions. The results clearly show that science teachers believe that boys and girls like different kinds of subjects. Some of the differences that they perceive between the preferences of boys and girls are only quantitative differences. For example, although both sexes prefer subjects that are practical, factual, creative, simple and important, boys are more attracted than girls to subjects that are practical, factual and important, whilst girls are more attracted than boys to subjects that are creative and simple. Other perceived differences are of a qualitative nature. Thus boys are thought to prefer subjects that can be described as numerical, science, logical and masculine. In contrast, the teachers believe that girls prefer subjects that are verbal, arts, intuitive and feminine.

7.2.2.2 Science teachers' responses compared with other teachers' responses

Comparing the ratings reported above, which were made by the science teachers in the main study, with ratings made by other groups of secondary teachers, particularly non-science teachers, there appears to be much agreement among secondary teachers over the characteristics that boys and girls prefer subjects to display (see Appendix 7.3, Table A7.3/1). Evidence that secondary school teachers are more stereotyped in their beliefs regarding pupils' preferences for subject characteristics than are middle school or primary school teachers is also presented in Appendix 7.3 (TableA7.3/2).

Table 7.15 Factor analysis of girls perceived preference for subject characteristics: Varimax factor loadings

Variable	Rotated factors 1	2	3	4
Practical-theoretical	46*	-07	0	-14
Numerical-verbal	-48*	16	37	36
Science-arts	19	02	71*	38
Logical-intuitive	-04	62*	10	07
Masculine-feminine	-11	06	34*	-02
Factual-opinionative	0	60*	03	04
Routine-creative	-43*	23	24	-33
Complex-simple	-02	09	11	53*
Important-unimportant	63*	18	0	18

Decimal points omitted
Salients asterisked

Table 7.16 Factor analysis of boys perceived preference for subject characteristics: Varimax factor loadings

Variable	Rotated factors 1	2	3	4
Practical-theoretical	08	44*	-08	-03
Numerical-verbal	29	05	59*	01
Science-arts	83*	30	11	04
Logical-intuitive	79*	20	06	-07
Masculine-feminine	07	54*	04	08
Factual-opinionative	39	59*	17	10
Routine-creative	01	-02	-02	59*
Complex-simple	-17	-02	20*	-03
Important-unimportant	13	43*	03	-17

Decimal points omitted
Salients asterisked

7.2.2.3 Science teachers' responses compared with pupils' responses

A simple way of measuring the accuracy of science teachers' beliefs regarding pupils' preferences for subject characteristics is to compare the responses of the teachers with pupils' responses to the same question. Table A7.3/3 in Appendix 7.3 presents a selection of results obtained from both teachers and pupils. They indicate that girls' preferences for subject characteristics are closer to the characteristics that boys prefer than teachers realize. The discrepancies between teachers' opinions about boys' preferences for different subject characteristics and boys' actual preferences are more variable. The teachers overrated the attraction of some characteristics and underrated the attraction of others. Since teachers tended to give more extreme ratings to girls' preferences and boys' preferences than the pupils did themselves, this means that the teachers believe that there are greater differences between boys' preferences and girls' preferences than there actually are.

7.2.2.4 Factor analysis of science teachers' responses

Science teachers' replies to the two scales were factor analysed separately to gain insight into the dimensions underlying teachers' perceptions of the subject characteristics preferred by boys and by girls. The technique of factor analysis is discussed further in Appendix 6.7.

Analysis of girls' preferences produced four factors with eigenvalues greater than 1. These four factors accounted for 65.3% of the total variance. They were subsequently rotated using the Varimax method. The factor loadings obtained are recorded in Table 7.15. Likewise, analysis of boys' preferences produced four factors with eigenvalues greater than 1. These factors, which accounted for 67.0% of the total variance, were also rotated. Table 7.16 records the factor loadings obtained.

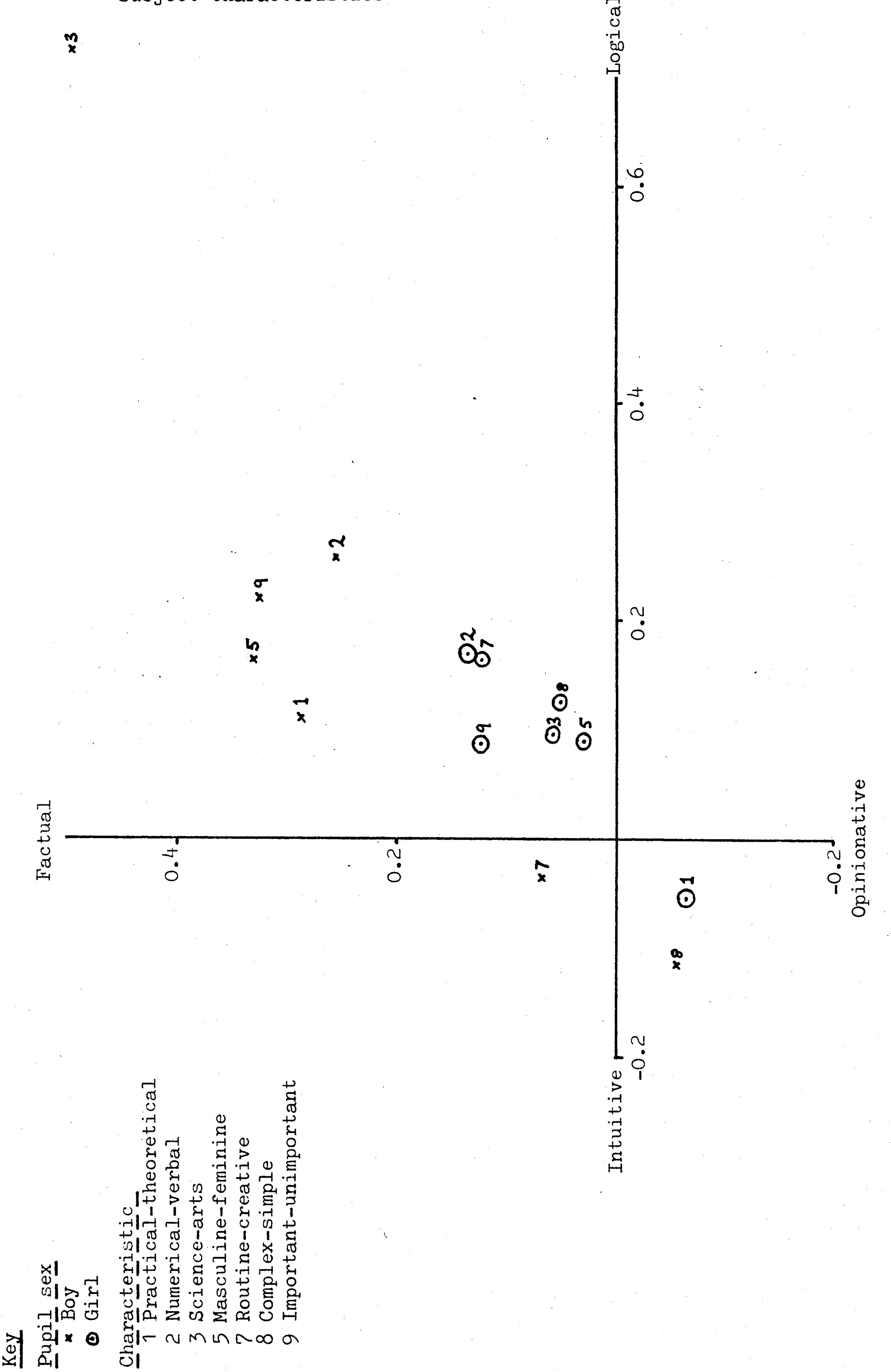
The most striking point to emerge from Tables 7.15 and 7.16 is that whereas the variables masculine/feminine and science/arts fall within the same factor for girls' preferences, they are found in different factors for boys' preferences. This suggests that science teachers, when describing subject characteristics preferred by girls, associate the two characteristics 'feminine' and 'arts' more closely than they do the two characteristics 'masculine' and 'science' when describing subject characteristics preferred by boys. Thus it seems that science teachers tend to believe that girls prefer subjects that are both arts and feminine. On the other hand, they tend to believe that boys prefer subjects that are masculine and also subjects to which the label 'science' can be attached.

The position of the variable routine/creative in the factor structure of boys' preferences is curious. Presumably, the teachers either regarded it as an inappropriate variable, or else they genuinely believe that it is unrelated to the other variables. However for girls' preferences, creativity in a subject is clearly linked with it being a verbal subject. These observations suggest that science teachers believe that creative subjects have different associations for boys and for girls.

The above factor analyses were conducted on replies from 159 respondents. Because of the small sample size (see Appendix 6.7 for the minimum recommended size), the data were further analysed using an alternative technique. Youngman's correlation scatter plotting method (1981) was chosen to provide a quick representation of the variable intercorrelations. This method can produce results that correspond very closely with those generated by a varimax factor analytic solution (Youngman, 1981).

Two variables were selected to be used as orthogonal reference axes. Ideally these variables needed to exhibit a low correlation with each other, but exhibit high correlations with many of the other variables in the correlation matrix. The variables 'logical-intuitive' and 'factual-opinionative' met these two criteria. A scatter plot was then produced

Figure 7.5 Scatter plot of pupils' perceived preferences for subject characteristics



by locating the remaining variables at points corresponding to the pair of correlations between each variable and the two reference variables.

Figure 7.5 shows the resulting scatter plots for girls' preferences and boys' preferences. Several clusters of variables are immediately obvious. The tightest cluster refers to girls' preferences and contains three variables - arts, feminine and simple. This cluster corresponds to Factor 3, and corroborates the close link that science teachers make between girls liking for subjects that are both feminine and arts subjects. The second cluster referring to girls' preferences contains another three variables - verbal, creative and important. This cluster corresponds to Factor 1. The most readily identifiable cluster referring to boys' preferences consists of three variables - practical, masculine and important. The variable 'science' is set apart from this cluster. The cluster corresponds to Factor 2, and 'science' represents Factor 1, 'logical' not having been plotted since the variable forms one of the axes. Again, the scatter plot corroborates the factor analysis finding that the variables 'science' and 'masculine' are not closely interrelated when applied to boys' preferences for subject characteristics.

7.2.2.5 Summary

1. Science teachers believe that boys' and girls' preferences for subject characteristics differ significantly. Most noticeably, boys are thought to prefer subjects that can be described as numerical, science, logical and masculine; whereas girls are thought to prefer subjects that are verbal, arts, intuitive and feminine.

2. A comparison between the responses from groups of science teachers and pupils revealed that teachers believe that boys' and girls' preferences for subject characteristics are more dissimilar than they actually are.

3. Factor analysis suggests that science teachers link the two variables 'arts' and 'feminine' together closely when describing subject

characteristics preferred by girls, but do not make a similar linkage between the variables 'science' and 'masculine' when describing subject characteristics preferred by boys. A correlation scatter plot corroborated this finding.

7.2.3 Females' Social Roles

The results reported below largely arise from the 5-item Attitudes to Females' Social Roles scale used in the main study. However, reference is also made to data gathered during the two pilot studies (see Appendix 6.11). The pilot findings provide additional support for the trends detected in the main study. Furthermore, since very much longer scales were used in the pilot studies, information was collected about a wide range of beliefs and opinions. This information helps to place the sentiments expressed in the 5-item scale into a broader context.

7.2.3.1 Pilot studies

Briefly summarizing the attitudes endorsed in the pilot studies, it was found that only 51% of the science teachers agreed that women are as good as men at complicated technical matters, whilst 13% disputed that women are men's equals intellectually. 16% agreed that women are not suited to jobs of great stress and responsibility. Furthermore, considerable doubt was expressed about the importance of women's careers. 42% agreed that a woman's career is not as important as a man's, and 37% thought that men will always be the basic breadwinners. There was more than a slight feeling that women should be occupied with domestic activities. 18% agreed that women's most important job is to look after the comforts of men and children, and 29% thought that a woman's place is in the home. Looking after children is viewed as an important task for women. 76% agreed that a good mother would not go out to work whilst she had a child under 5. The last area explored by the attitude items concerned relationships between men and women, and in particular between

Table 7.17 Breakdown of science teachers' total scores on the Females' Social Roles Questionnaire (Main study, N=163)

Total score	Percentage
5 - 9	6.1
10 - 14	24.6
15	9.8
16 - 20	35.7
21 - 25	23.8

Table 7.18 Females' Social Roles scale items and science teachers' responses (Main study, N=163)

	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
1. As head of the household, the father should have final authority over his children.	9.2	13.5	12.3	33.1	31.9
2. A woman who refuses to give up her job to move with her husband would be to blame if the marriage broke up.	3.7	22.7	27.0	28.2	18.4
3. A woman who refuses to bear children has failed in her duty to her husband.	8.0	9.8	18.4	28.2	35.6
4. A woman should be expected to change her name when she marries.	8.0	28.2	31.1	21.5	11.0
5. It is all right for women to work but men will always be the basic breadwinners.	9.8	24.5	11.7	26.4	27.6

husbands and wives. A sizeable minority of the respondents indicated that women should be subordinate to men. 29% thought that it is the man's job to make the major decisions. 20% agreed that women should obey their husbands, and 16% agreed that a woman should allow her husband to feel superior even if this involves belittling herself.

7.2.3.2 Main study

Table 7.17 presents a breakdown of the total scores awarded in the main study. The lowest total score possible was 5, and the highest was 25, with a mid-score of 15. It can be seen that 40% of the teachers obtained a score of 15 or less, indicating that they gave traditional, sex stereotyped replies. Of the 60% of teachers who gave more liberal replies, 24% actually gave very liberal replies. These very high total scores were mostly achieved by female teachers. The mean total score for female teachers was 19.6, as opposed to 15.9 for male teachers. This difference is highly significant ($t = 5.30$, $p < 0.001$).

A detailed analysis of the replies to the individual items is presented in Table 7.18. The figures show that on average about a quarter of the respondents agreed with each statement. It is particularly interesting to note that as many as a third of the respondents thought that men will always be the basic breadwinners.

Replies to individual items were analysed further by comparing the answers given by known groups of teachers. First the replies of men and women were compared (Table 7.19). As hypothesized, it was found that women gave significantly more liberal replies than men on all five items (one-tailed t test). As mentioned previously, the total scores of women were also significantly more liberal than those of men. The same result emerged from both pilot studies as well. Women obtained significantly more liberal total scores than did the men (see Appendix 7.5). Another teacher variable that produced interesting results was that of age. As hypothesized, older teachers gave more traditional replies than younger

Table 7.19 Females' Social Roles scale: Item means of male and female teachers

Item No.	Teacher sex	$\bar{x}$	s.d.	t	p
1	Male	3.34	1.36	-5.47	.001
	Female	4.30	0.87		
2	Male	3.24	1.16	-1.86	.05
	Female	3.58	1.05		
3	Male	3.46	1.30	-4.61	.001
	Female	4.30	0.97		
4	Male	2.86	1.10	-2.15	.001
	Female	3.26	1.13		
5	Male	3.02	1.33	-5.14	.001
	Female	4.11	1.14		
N	Male	110			
	Female	53			

Table 7.20 Females' Social Roles scale: Mean scores of different age groups

Age	N	Mean score
Under 30	29	18.21
30 - 39	73	17.33
40 - 49	30	16.37
50 & over	13	15.46

teachers (see Table 7.20). When the teachers were split into two age groups, above and below 40, then the older age group obtained significantly lower scores, i.e. more traditional scores, than did the younger age group (see Appendix 7.5). None of the other comparisons investigated consistently produced significant differences. However, certain trends were detected and these are recorded in Appendix 7.5.

7.2.3.3 Summary

1. A sizeable minority of science teachers hold very traditional, sex stereotyped attitudes about women's social roles. The two pilot studies inquired into a wide range of topics and the following points emerged.

(a) A small proportion of teachers believe women's intellectual abilities to be inferior to those of men. Doubts about women's technical competence are much more widespread.

(b) Considerable doubt was expressed about the importance of women's careers and their significance as wage earners.

(c) Nearly one third of respondents agreed that a woman's place is in the home. Childcare was viewed as the sole responsibility to the mother by the majority of respondents.

(d) A sizeable minority of the respondents indicated that women should be subordinate to men.

2. The short-form scale used in the main study established that approximately one quarter of the respondents agreed that:

(a) women's careers and earning capacity are secondary to those of men,

(b) the man is the dominant and controlling partner in a marriage.

The total scores indicated that 40% of the science teachers held traditional sex role beliefs.

3. Overall, women gave significantly more liberal replies than men, and younger teachers gave more liberal replies than older teachers.

Table 7.21 The importance of different subject areas to pupils' general education as judged by secondary teachers of all subjects (N=67)

Subject Area	Mean rating Boys	Girls	d	t	p
Creative Arts	3.09	3.25	-0.24	-3.28	0.01
Languages	3.25	3.33	-0.11	-1.92	ns
Humanities	3.37	3.43	-0.10	-1.42	ns
Science	3.82	3.58	0.50	4.51	0.001
Technical Subjects	3.48	2.91	0.85	6.83	0.001
Home Economics	2.54	3.16	-0.82	-7.21	0.001
Commercial/Business Studies	2.60	2.91	-0.39	-3.44	0.01

Table 7.22 The importance of qualifications in different subject areas to pupils' future lives as judged by secondary teachers of all subjects (N=35)

Subject Area	Mean rating Boys	Girls	d	t	p
Creative Arts	3.20	3.23	-0.04	-0.57	ns
Languages	3.23	3.23	0.0	0.0	ns
Humanities	3.26	3.23	0.04	0.98	ns
Science	3.71	3.54	0.32	2.64	0.05
Technical Subjects	3.60	2.83	1.10	5.63	0.001
Home Economics	2.74	3.17	-0.46	-3.63	0.01
Commercial/Business Studies	3.11	3.34	-0.29	-2.49	0.05

Table 7.23 The importance of qualifications in different subject areas to pupils' future lives as judged by secondary science teachers (N=62)

Subject Area	Mean rating Boys	Girls	d	t	p
Creative Arts	2.50	2.84	-0.45	-3.69	0.001
Languages	2.77	2.85	-0.10	-1.52	ns
Humanities	2.68	2.92	-0.35	-3.79	0.001
Science	3.65	3.23	0.70	6.64	0.001
Technical Subjects	3.61	2.60	1.47	9.53	0.001
Home Economics	2.32	3.06	-1.01	-7.84	0.001
Commercial/Business Studies	2.77	3.08	-0.41	-3.09	0.01

7.2.4 Importance of Subjects

The replies which had been collected using a 4-point verbal rating scale, were converted to numerical data by assigning the value 1 to 'not at all important' and so on up to 4 for 'very important'. Thus a high rating denotes that teachers believe that that subject area is important in a pupil's education, whilst a low rating implies that the subject area is of little importance.

The mean ratings obtained for boys and girls in each of the different subject areas on each administration are recorded in Tables 7.21 to 7.23. Tables 7.21 and 7.22 actually refer to results from the two pilot studies. These results are included since they provide valuable comparative and confirmative information. Paired t tests were used to investigate the statistical significance of the difference between the mean ratings given for boys and for girls. The t values obtained are shown in the tables, together with indications of their significance under two-tailed test conditions. Statistical significance does not necessarily imply educational significance, so d values (see Appendix 7.6) are also included in the tables since these figures offer a convenient measure of the magnitude of the differences.

7.2.4.1 The views of teachers in general

The results clearly show that teachers believe that some subject areas are of greater importance to one sex than to the other. The teachers indicated that science and technical subjects are of greater importance to boys' education, whereas home economics and commercial/business studies are more important to girls' education. These views were expressed whether teachers were asked to consider the general education of pupils (Table 7.21) or the more specific question of the value of qualifications in the different subject areas to pupils' future lives (Table 7.22).

Comparing the results presented in Tables 7.21 and 7.22,

Table 7.24 The representativeness of science teachers' views about the importance of qualifications in different subject areas to pupils' future lives

(A) Male pupils

Subject area	Mean ratings		z	p
	Science teachers	Teachers of all subjects		
Creative Arts	2.50	3.20	-4.49	0.001
Languages	2.77	3.23	-2.86	0.01
Humanities	2.68	3.26	-3.79	0.001
Science	3.65	3.71	-0.67	ns
Technical Subjects	3.61	3.60	0.12	ns
Home Economics	2.32	2.74	-2.23	0.05
Commercial/Business Studies	2.77	3.11	-2.00	0.05

(B) Female pupils

Subject area	Mean ratings		z	p
	Science teachers	Teachers of all subjects		
Creative Arts	2.84	3.23	-2.42	0.05
Languages	2.85	3.23	-2.42	0.05
Humanities	2.92	3.23	-1.99	-.05
Science	3.23	3.54	-2.42	-.05
Technical Subjects	2.60	2.83	-1.30	ns
Home Economics	3.06	3.17	-0.59	ns
Commercial/Business Studies	3.08	3.34	-1.68	ns

particularly the d values, reveals that the teachers expressed more sex differentiated views when they were asked a somewhat vague, general question. The responses received for technical subjects constitute the only exception to this tendency. When considering pupils' future lives, teachers are even more convinced that technical subjects are likely to be of greater importance to boys than to girls.

Although the replies received to the specific question were in general less sex differentiated than those received to the abstract question, the ratings given to boys and to girls separately for each subject area under the two conditions were not significantly different, with the exception of the ratings for commercial/business studies. Qualifications in commercial/business studies were judged to be significantly more important to both boys' future lives ($z = 3.06$, $p < 0.01$) and to girls' future lives ($z = 2.74$, $p < 0.01$) than was the contribution of the subject group to their general education.

7.2.4.2 Science teachers' responses compared with other teachers' responses

An inspection of the mean ratings received from teachers of various subjects and those received from science teachers (Tables 7.22 and 7.23), and a comparison of the d values for the two groups shows that the science teachers gave the more sex differentiated responses. They were more convinced of the greater value of science and technical subjects for boys, and home economics and commercial/business studies for girls. In addition, they also believed that qualifications in creative arts and humanities are more important for girls than boys.

Besides giving more sex differentiated responses, the science teachers also tended to place less importance on qualifications in the different subject areas under investigation (see Table 7.24). The importance of qualifications in the arts subjects, namely creative arts, languages and humanities, to both boys' and girls' future lives was rated significantly lower by science teachers than by teachers of various subjects. Home economics and commercial/business studies were also devalued as subjects for boys by the science teachers, but not as subjects for girls. Interestingly, science teachers regarded the acquisition of qualifications in the science subjects to be significantly less important for girls than did teachers of other subjects. The only subject group that was valued similarly by the group of science teachers and the group of mixed teachers was technical subjects.

7.2.4.3 Summary

1. Teachers of all subjects believe that science and technical subjects are of greater importance to boys' education than to girls' education, whereas home economics and commercial/business studies are more important to girls' education.

2. Science teachers gave the most sex differentiated replies. They believe more firmly in the greater value of science and technical subjects for boys, and the value of home economics, commercial/business studies, creative arts and humanities for girls.

3. With the exception of technical subjects, science teachers regarded qualifications in a range of optional subject areas to be less important to pupils' future lives than did teachers of other subjects.

4. Surprisingly, science teachers regarded the acquisition of qualifications in the science subjects to be significantly less important for girls than did teachers of other subjects.

7.2.5 Conclusions

1(a) 85% of science teachers were able to list features that they consider to be typical of the written work of each sex. Thus Hypothesis Four is confirmed. Science teachers recognise differences between the written work of girls and boys.

(b) (Relates to Question Eight, Do similar proportions of science teachers and other teachers recognise differences between the written work of girls and boys?) The proportion of teachers who listed features typical of the written work of each sex was very similar for both science teachers (85%) and teachers of other subjects (82%).

(c) (Relates to Question Six, Do science teachers believe that they can generally distinguish between the written work of girls and boys?) A majority of science teachers claimed that they can generally distinguish between the written work of each sex. Interestingly, a significantly higher proportion of male teachers (80%) than female teacher (57%) believe that they can identify the work of girls and boys.

(d) (Relates to Question Nine, Do similar proportions of science teachers and other teachers believe that they can distinguish between the written work of girls and boys?) The proportion of teachers who claimed that they can tell the difference between boys' and girls' written work was nearly identical for both science teachers (72%) and teachers of other subjects (73%).

(e) (Relates to Question Seven, What features do science teachers associate with the written work of girls and boys?) Science teachers describe girls as being more conscientious over their work. Consequently their written work is neater, better presented, more thorough and lengthier. Also their handwriting is more regular and easier to read. In contrast, boys are seen to be more careless over their work, and so the work is untidy, poorly presented and usually brief. However, the written work of boys is credited as indicating more understanding than that of girls.

(f) (Relates to Question Ten, Do science teachers and other teachers associate similar features with the written work of girls and of boys?) In general, science and non-science teachers refer to similar features when describing the written work of boys and girls (Table 7.13). Both use similar terms to describe the handwriting and the overall appearance of boys' and girls' work. However, science teachers are more aware of differences in the quantity of work produced by boys and girls, and also the quality of its content than are non-science teachers. The only characteristic that was mentioned by non-science teachers, but not science teachers, was the type of topic that boys and girls choose to write about.

2(a) Science teachers believe that boys' and girls' preferences for subject characteristics differ significantly (Table 7.14). Most noticeably, boys are thought to prefer subjects that can be described as numerical, science, logical and masculine; whereas girls are thought to prefer subjects that are verbal, arts, intuitive and feminine. These findings provide support for Hypothesis Five. Science teachers believe that boys and girls have different preferences regarding subject characteristics.

(b) (Relates to Question Eleven, How do secondary science teachers' perceptions of pupils' preferences for different subject characteristics compare with those of teachers of other subjects and teachers from other educational levels?) Secondary teachers, of both science and non-science subjects, are in broad agreement over the characteristics that boys and girls prefer subject to display (Table A7.3/1, Appendix 7.3). However, science teachers are marginally more certain that both boys and girls prefer science subjects. Compared to middle school and primary school teachers, there are indications that secondary school teachers (including science teachers) are more stereotyped in their beliefs regarding pupils' preferences for subject characteristics. For example, secondary teachers believe more firmly that girls prefer subjects that are simple and verbal. Their beliefs about the attraction of a subject's gender image are even more stereotyped. Not only are secondary teachers more extreme in their view that girls prefer feminine subjects, but they also believe more firmly that boys prefer masculine subjects (Table A7.3/2, Appendix 7.3).

3(a) On a scale measuring attitudes towards females' social roles, 40% of science teachers obtained low scores (half the maximum score or less) indicating that they gave traditional, sex stereotyped replies. An alternate way of stating the results would be to say that of the majority of science teachers who gave slightly liberal replies, 24% gave very liberal replies. However, the fact should not be overlooked that a very sizeable minority gave traditional replies. Thus Hypothesis Six received some support, but not unequivocal support. Some science teachers believe in traditional sex roles.

(b) Women obtained significantly more liberal scores on the sex role attitude scale than did men (Table 7.19). This finding confirms Hypothesis Seven. Male science teachers hold more traditional attitudes towards sex roles than do female science teachers.

(c) Significantly more liberal scores on the sex role attitude scale were obtained by teachers who were under 40 than were obtained by teachers who were 40 and over. This finding confirms Hypothesis Eight. Older science teachers hold more traditional attitudes towards sex roles than do younger science teachers.

4(a) Science teachers do not regard qualifications in all subject areas to be equally important for boys' and girls' future lives. They believe that science and technical subjects are of greater importance to boys. Creative arts, humanities, home economics and commercial/business studies are judged to be of greater value to girls (Table 7.23). These findings support Hypothesis Nine. Science teachers believe that not all school subjects are of equal importance for boys and for girls.

(b) (Relates to Question Fourteen, Do science teachers consider the science subjects to be of greater value and relevance to boys than to girls?) Not only do science teachers regard the acquisition of qualifications in the science subjects to be significantly less important for girls than for boys (Table 7.23), but they even judge science to be of less importance to girls than do teachers of other subjects (Table 7.24).

Table 7.25 Ranking and mean rating of factors believed to contribute to pupils' success in science (N=163)

Rank	Variable	$\bar{X}$	SD
1	Effort	1.23	0.42
2	Good attitude	1.33	0.50
3	Ability	1.41	0.52
4=	Interest in subject	1.50	0.53
4=	Conscientiousness	1.50	0.54
6	Attentiveness	1.52	0.54
7	Good teaching	1.63	0.58
8	Family support	2.12	0.75
9	Emotional stability	2.14	0.68
10	Out-of-class experience	2.53	0.70
11	Subject simplicity	2.56	0.69
12	Assistance from peers	2.58	0.66

7.3 ATTRIBUTION PATTERNS

7.3.1 Reasons for Success at Science

The replies, which had been collected using a 4-point verbal rating scale, were converted to numerical data by assigning the value 1 to 'very important', 2 to 'fairly important', 3 to 'not very important' and 4 to 'not at all important'. Thus low numbers denote that a variable is considered to be an important one.

7.3.1.1 Relative importance of variables

Firstly, the teachers views about the comparative importance of each factor to success was investigated. By averaging all the replies, an indication of the comparative importance of each variable to pupils' success in the science subjects generally was obtained. The mean ratings were then ranked to show more clearly those factors that the teachers considered were most important to success in science, and those factors believed to contribute least to success in science. Table 7.25 shows the mean ratings given to each factor and their rank position.

The first six factors appearing in Table 7.25 all refer to attributes or behaviours displayed by the pupil, i.e. internal causes. Of the external factors, good teaching heads the list, followed by family support. Out-of-class experience, subject simplicity and assistance from peers were all deemed to contribute very little to success in science.

7.3.1.2 Effect of pupil sex

Since the main object of the scale was to detect differences between the patterns of factors used to explain boys' and girls' success in science, attention was next directed to the responses given for boys and girls separately. To avoid confounding differences in replies due to subject characteristics with those due to pupil sex, the replies were analysed for each science subject separately. Mean ratings for boys and girls in each of the three main science subjects were plotted for each

Table 7.26 The contributory factors to success in science for boys and girls, clustered by elementary linkage analysis

Boys (N=73)	Girls (N=90)
1. Effort	1. Ability
Conscientiousness	Good attitude
Attentiveness	Interest in subject
	Good teaching
2. Out-of-class experience	Effort
Family support	Conscientiousness
	Attentiveness
3. Good teaching	
Emotional stability	2. Out-of-class experience
	Emotional stability
4. Ability	Family support
Good attitude	
Interest in subject	3. Subject simplicity
Subject simplicity	Assistance from peers
Assistance from peers	

variable. The graphs indicated slight differences in the attribution patterns for the different subjects, e.g. out-of-class experience was considered to contribute much less to success in chemistry than to success in physics or biology. The attribution patterns for boys and girls in each subject were generally very similar, with only minor differences, e.g. family support was judged to be more important to girls' than boys' success in chemistry, whereas the reverse was the case for biology. To ascertain whether any of the differences between the mean ratings associated with boys' and girls' success were statistically significant, two-tailed t tests were calculated. None of them was significant, with the exception of the contribution of effort to success in physics. Pupil effort was believed to be more important to girls' success in physics than to boys' success ($t = 2.62$, $p < 0.05$).

7.3.1.3 Cluster analysis

Although no obvious differences were detected in the teachers' attribution patterns for boys' and girls' success in the different science subjects, it was still possible that the teachers were grouping the factors differently for boys and for girls. To investigate the inter-relationships between the different factors and the underlying dimensions, some form of cluster analysis or factor analysis was required. Elementary linkage analysis (Appendix 6.6) was chosen in preference to factor analysis because the sample sizes were rather small (see discussion in Appendix 6.7). Furthermore, because of the small size of the samples (N=73 (Boys), N=90 (Girls)), it was thought expedient to pool the data and consider science in general, rather than the separate subjects.

The 12 factors under investigation, that are believed to contribute to success, were grouped into four clusters when used to explain boys' success in science. However, only three clusters emerged when they were used to account for the success of girls in science. Table 7.26 shows the factors that appeared in each cluster for boys and for girls.

Table 7.27 Teachers' mean ratings of the contribution of ability to pupils' success in the science subjects (N=147)

	Pupil sex	
Subject	Boy	Girl
Physics	1.38	1.64
Chemistry	1.35	1.21
Biology	1.30	1.43

Table 7.28 Teachers' mean ratings of the contribution of effort to pupils' success in the science subjects (N=147)

	Pupil sex	
Subject	Boy	Girl
Physics	1.50	1.12
Chemistry	1.23	1.21
Biology	1.30	1.14

A detailed inspection of the composition of the clusters recorded in Table 7.26 reveals that for both boys and girls the first cluster refers mainly to factors internal to the pupil, whilst the second cluster refers to factors that are mainly external to the pupil. Clusters 3 and 4 for boys contain a mixture of factors, but the third cluster for girls contains external factors. Thus two of the three clusters formed by the attributions for girls contain external factors, whilst only one of the four boys' clusters consists solely of external factors. All the important internal factors are found in a single cluster, the first cluster, for girls but they are more dispersed for boys.

7.3.1.4 Analysis of variance

More complex interrelationships between pupil sex and other independent variables were investigated by analysis of variance (see Appendix 7.7). A 2x2x3 (sex of pupil, sex of teacher, principal teaching subject of teacher) analysis of variance was computed for each causal factor.

As prior analysis had indicated, no main effects for pupil sex emerged. Principal teaching subject produced one main effect. Out-of-class experience was considered to be less important to success in chemistry than to success in physics and biology, $F(2,136) = 7.30$, $p < 0.001$. Two main effects of teacher sex showed that female teachers considered effort and emotional stability to be more important contributory factors to success in science than did male teachers, $F(1,136) = 4.89$, $p < 0.05$, and $F(1,136) = 5.92$, $p < 0.05$ respectively.

Two interesting interaction effects between pupil sex and teaching subject emerged (see Tables 7.27 and 7.28). Ability was judged to contribute more to boys' success in physics and biology, but not in chemistry, $F(2,136) = 3.56$, $p < 0.05$. In contrast, effort was considered to contribute more to girls' success in all three subjects. Effort was judged to be most important to girls' success in physics and least important to boys' success in physics, $F(2,136) = 3.32$, $p < 0.05$. Pupil sex and teacher sex produced one second-order interaction for family support. Besides the fact that female teachers thought that family support contributed more to success in science than did male teachers, both female and male teachers indicated that family support was more important for boys than for girls, $F(1,136) = 4.39$, $p < 0.05$. The causal factor of family support also produced a third-order interaction, $F(2,136) = 4.57$, $p < 0.05$.

7.3.1.5 Summary

1. Science teachers believe that internal factors, such as a pupil's effort, ability and interest in the subject, contribute more to success in science than do external factors, such as family support, out-of-class experience.

2. Similar attribution patterns were used by science teachers to explain the success of both boys and girls in each of the three main science subjects, with one notable discrepancy. Pupil effort was judged to contribute significantly more to girls' success in physics than to boys' success.

3. Cluster analysis indicated that internal factors were dispersed through most of the different dimensions underlying the teachers' attributions for boys' success, whereas the attribution dimensions underlying girls' success were strewn with external factors.

4. Analysis of variance revealed that not only do female teachers think that family support contributes more to success in science than do male teachers, but that both female and male teachers believe that family support is more important for boys than for girls.

5. In the factor analysis, pupil sex and teaching subject produced two interesting second-order interations for the variables 'ability' and 'effort'.

Table 7.29 Ranking and mean rating of factors believed to contribute to pupils' failure in science (N=163)

Rank	Variable	$\bar{X}$	SD
1	Lack of effort	1.22	0.43
2	Poor attitude	1.31	0.50
3	Lack of interest in science	1.49	0.54
4	Lack of attention	1.60	0.52
5	Lack of ability	1.61	0.61
6	Subject difficulty	1.91	0.66
7	Poor teaching	1.99	0.69
8	Carelessness	2.01	0.64
9	Distraction by peers	2.04	0.71
10	Lack of emotional stability	2.22	0.75
11	Lack of family support	2.33	0.76
12	Little relevant out-of-class experience	2.64	0.81

7.3.2 Reasons for Failure at Science

Verbal replies were converted to numerical data as described in section 7.3.1. Thus the lower the number assigned to a variable, the more important that variable was judged to be.

7.3.2.1 Relative importance of variables

The replies for each factor were averaged and then the mean ratings were ranked in order of importance to show clearly those factors that the teachers thought contributed most to failure in science, and those factors believed to contribute least to failure in science (see Table 7.29).

The first five factors appearing in Table 7.29 all refer to personal attributes or behaviours displayed by the pupil, i.e. internal causes. Of the external factors, subject difficulty heads the list, followed by poor teaching. Lack of family support and little relevant out-of-class experience were thought to contribute very little to failure in science.

7.3.2.2 Effect of pupil sex

As a preliminary stage in looking for differences between the attribution patterns used to explain boys' and girls' failure in science, the mean ratings given to boys and girls for each variable were plotted for each of the three science subjects separately. The graphs indicated slight differences in the attribution patterns for the different subjects, e.g. distraction by peers was considered to contribute more to failure in biology than to failure in physics. The attribution patterns for boys and girls in each subject were generally very similar. One of the more obvious differences occurred in physics. Subject difficulty was judged to contribute more to girls' failure in physics than to boys' failure. Two-tailed t tests were calculated to determine whether any of the differences between the mean ratings associated with boys' and girls' failure were statistically significant. Only two subject-factor

Table 7.30 The contributory factors to failure in science for boys and girls, clustered by elementary linkage analysis

Boys (N=73)	Girls (N=90)
1. Poor teaching Poor attitude Lack of effort	1. Lack of interest in subject Poor attitude Lack of effort
2. Lack of interest is subject Lack of attention	2. Lack of family support Lack of emotional stability Little relevant out-of-class experience Poor teaching
3. Little relevant out-of-class experience Carelessness Distraction by peers Lack of family support Lack of emotional stability	3. Lack of ability Subject difficulty
4. Lack of ability Subject difficulty	4. Carelessness Lack of attention Distraction by peers

combinations gave significant t values ($p<0.05$). Subject difficulty was believed to contribute more to girls' failure in physics, and lack of relevant out-of-class experience to boys' failure in biology.

7.3.2.3 Cluster analysis

Next elementary linkage analysis was preformed to detect inter-relationships between the different factors and underlying dimensions. The 12 factors fell into four similar clusters when used to explain both boys' and girls' failure in science (Table 7.30).

7.3.2.4 Analysis of variance

The data were further analysed by analysis of variance. A 2x2x3 analysis (sex of pupil, sex of teacher, principal teaching subject of teacher) was computed for each causal factor. Principal teaching subject produced three main effects. Lack of ability was thought to contribute to failure least in physics and most in chemistry, $F(2,136) = 3.35$, $p < 0.05$. Likewise, subject difficulty was judged to be most important in chemistry, but least important in biology, $F(2,136) = 3.64$, $p < 0.05$. Regarding the influence of distraction by peers, this was considered to be a much more important factor contributing to failure in biology than in physics or chemistry, $F(2,136) = 7.04$, $p < 0.001$.

Teacher sex and pupil sex each produced one main effect and in both cases it referred to the causal factor 'lack of family support'. Female teachers considered lack of family support to be a more important contributory factor to failure in science than did male teachers, $F(1,136) = 2.26$, $p < 0.05$. Regarding pupil sex, lack of family support was believed to contribute more to boys' failure than to girls' failure, $F(1,136) = 5.54$, $p < 0.05$. Teacher sex and pupil sex also interacted to produce a second-order effect for lack of family support, $F(1,136) = 4.80$, $p < 0.05$. The only other second-order effect that emerged was another teacher sex and pupil sex interaction for the causal factor of distraction by peers. The interaction arose because distraction by peers was judged to be a more important contributory factor to boys' failure than to girls' failure, and because male teachers considered the factor to be more important to failure in science than did female teachers, $F(1,136) = 14.81$, $p < 0.001$. Finally, there was a third-order interaction effect for lack of relevant out-of-class experience, $F(2,136) = 4.71$, $p < 0.05$.

7.3.2.5 Summary

1. Science teachers believe that internal factors, such as lack of effort, poor attitude and lack of interest in the subject, contribute most to failure in science. Even though external factors are thought to be less influential, it is interesting to note that subject difficulty and poor teaching head the list of contributory external factors.

2. Similar attribution patterns were used by science teachers to explain the failure of both boys and girls in each of the three main science subjects, except that subject difficulty was believed to contribute more to girls' failure in physics, and lack of relevant out-of-class experience to boys' failure in biology.

3. External factors accounted for most of the main effects and all of the interaction effects that emerged from an analyis of variance. The interaction effects all involved pupil sex and affected three variables - lack of family support, distraction by peers and lack of relevant out-of-class experience.

7.3.3 Reasons for Choosing Science

The replies, which had been collected using a 5-point verbal rating scale, were converted to numerical data by assigning the value 1 to 'never applies', 2 to 'rarely applies', 3 to 'sometimes applies', 4 to 'often applies' and 5 to 'always applies'. Thus high numbers denote that a variable is thought to frequently influence subject choice.

Table 7.31 Ranking and mean rating of factors which influence pupils to choose science (N=164)

Rank	Variable	$\bar{X}$	SD
1	Like subject content	3.88	0.66
2	Relevancy for future career	3.76	0.67
3	Like teaching style	3.47	0.76
4=	Like teacher	3.31	0.74
4=	Parental influence	3.31	0.68
6	Find subject easy	3.29	0.78
7	Expect to like teaching group	3.19	0.89
8	Tradition	2.69	1.02
9	Boys'/girls' subject	2.64	1.16
10	Relevancy for future family life	2.37	0.91

7.3.3.1 Relative importance of variables

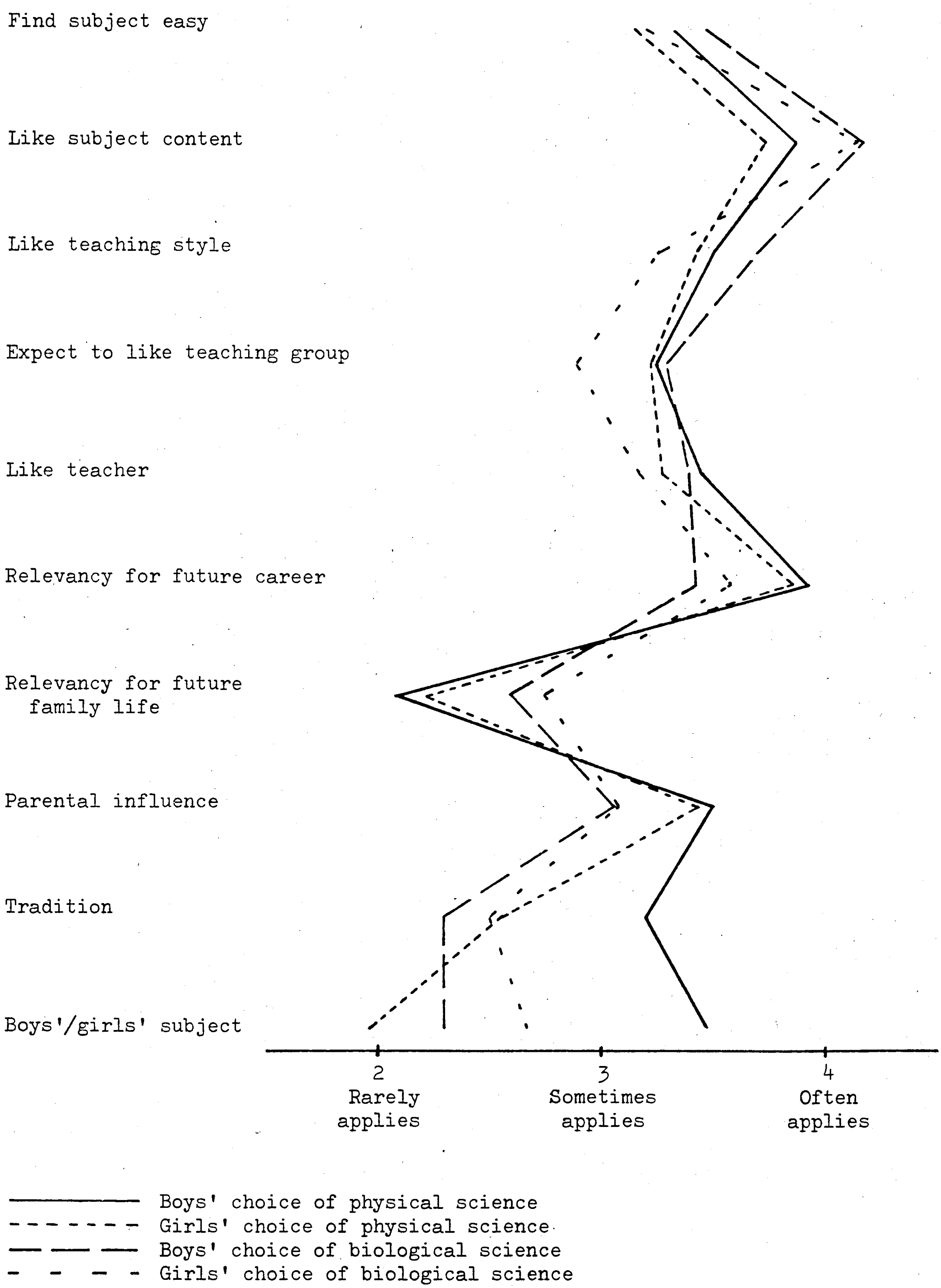
Firstly, the teachers' beliefs about the comparative influence of each factor upon subject choice was investigated. By averaging all the replies, an indication of the comparative influence of each variable upon pupils' choice of science subjects generally was obtained. The mean rating were then ranked to show more clearly those factors that the teachers considered were most important in causing pupils to choose science, and those factors believed to influence the choice of science least. Table 7.31 shows the mean ratings given to each factor and their rank position.

Table 7.31 shows that teachers believe that pupils mostly choose science subjects because they expect to like the topics covered in the subject, and because the subject is needed for their chosen careers. These two most frequently operating reasons are closely followed by teacher effects and parental influence. Thus teachers believe that most pupils' choice of science results from a mixture of school based factors and out-of-school pressures. Teachers do not consider that diffuse social factors, such as tradition and sex-typing of a school subject, are very important influences upon pupils' decision to choose science subjects. The least important reason, according to science teachers, why pupils choose science is that a study of science will help them in their everyday lives with their future families.

7.3.3.2 Effect of pupil sex

Since one of the main objects of the scale was to detect differences between the patterns of motivations used to account for boys' and girls' choice of science, attention was next directed to the responses given for boys and girls separately. To avoid confounding differences in replies due to subject characteristics with those due to pupil sex, the replies were analysed for the physical sciences and biological science separately. Mean ratings for boys and girls in each subject area were

Figure 7.6 Pupils' reasons for choosing science subjects as viewed by science teachers



plotted for each variable (Figure 7.6).

The graphs highlight the finding from the previous analysis that not all the reasons are equally compelling. Furthermore, they show that not all the reasons are equally appropriate for both science areas. Thus liking of the subject content and relevancy for future family life are thought to operate more frequently in persuading pupils to choose biology than physics and chemistry. On the other hand, pupils are more likely to choose a physical science subject than biology because of it relevance for future careers.

Figure 7.6 clearly displays the effect of pupil sex within each subject area. The graphs showing the frequency with which the factors influence boys and girls to choose physical science subjects are very similar, with the exception of large differences for tradition and boys' subject. Both of these differences are statistically significant at the 0.1% level (two-tailed t test). Although none of the values associated with the other factors are statistically different for boys and girls, it is worth noting that every variable was judged to be more important in influencing boys' decisions, with the exception of the subject's relevance for future family life, which happens to be the least influential variable anyway. Turning to biological science, the factors were equally important in influencing the choice of boys and girls. However, a closer inspection reveals that variables referring to characteristics of the subject and the teaching of the subject are believed to influence boys' choice more than girls' choice, whilst social and career considerations are more effective in persuading girls to take biology. Only one factor, teaching style, was judged to differ significantly in its effect upon boys' and girls' choice of biology ($t = 2.26$, $p < 0.05$).

7.3.3.3 Interaction between pupil sex and science subject

Interactions between pupil sex and science subject area can also be

Table 7.32 Pupils' reasons for choosing science subjects as viewed by science teachers

Reason	Pupil sex	Phys	Chem	Biol	Analysis of variance Significance: Pupil sex	Subject	Inter-action
Find subject easy	Boys	3.38	3.31	3.48	-	-	-
	Girls	3.12	3.20	3.21			
Like subject content	Boys	4.13	3.73	4.17	-	xxx	-
	Girls	3.76	3.70	4.14			
Like teaching style	Boys	3.75	3.35	3.70	-	-	-
	Girls	3.56	3.30	3.25			
Expect to like teaching group	Boys	3.63	3.04	3.30	-	-	-
	Girls	3.16	3.27	2.89			
Like teacher	Boys	3.63	3.35	3.39	-	-	-
	Girls	3.16	3.37	3.18			
Relevancy for future career	Boys	4.06	3.85	3.43	-	xx	-
	Girls	3.76	3.90	3.61			
Relevancy for fut. family life	Boys	2.38	1.88	2.61	-	xxx	-
	Girls	2.36	2.10	2.75			
Parental influence	Boys	3.69	3.38	3.04	-	xx	-
	Girls	3.36	3.47	3.07			
Tradition	Boys	3.69	2.88	2.30	x	x	xxx
	Girls	2.32	2.77	2.50			
Boys'/girls' subject	Boys	3.88	3.23	2.30	xxx	-	xxx
	Girls	1.88	2.00	2.68			
N	Boys	16	26	23			
	Girls	25	30	28			

xxx Significant at 0.1% level
xx Significant at 1% level
x Significant at 5% level
- Not significant

discerned in Figure 7.6. Some factors differ in their influence over boys and girls according to the type of science subject under consideration. For example, tradition is a much more powerful influence upon boys' choice of physical science than upon girls' choice of physical science, whilst the reverse is true for biological science.

Interactions between pupil sex and science subject were investigated further using analysis of variance. It was first established that male and female teachers did not differ significantly in their judgements regarding the importance of the different factors. Thus teacher sex was not included as an independent variable in the analysis. A 2x3 analysis (sex of pupil, science subject) was computed for each motivating reason. The results are recorded in Table 7.32.

Pupil sex produced two main effects. Tradition and the many associations linking science with males were both considered to influence boys more frequently than girls in their choice of science. Science subjects under consideration produced five main effects. The teachers believed that pupils are more likely to choose physics and chemistry than biology because of career considerations, parental influence and tradition. Biology is more often chosen for its relevance to family life and because pupils like the subject content. Chemistry is least frequently chosen for the two foregoing reasons. The two interaction effects for tradition and boys' subject are largely due to the teachers perception that boys often choose physics because it is a boys' subject and because boys have traditionally chosen to study physics and be associated with the subject.

7.3.3.4 Cluster analysis

Elementary linkage analysis (Appendix 6.6) was performed to detect interrelationships between the different factors and underlying dimensions. Because of the small size of the samples (N=73 (Boys), N=91 (Girls)), it was thought expedient to pool the data and consider science

Table 7.33 Reasons why boys and girls choose science, clustered by elementary linkage analysis.

Boys (N=73)	Girls (N=91)
1. Expect to like teaching group	1. Relevancy for future family life
Like teacher	Parental influence
Like teaching style	Tradition
Find subject easy	Boys'/girls' subject
Like subject content	
Relevancy for future family life	2. Expect to like teaching group
	Like teacher
2. Relevancy for future career	Like teaching style
Parental influence	Find subject easy
Tradition	Like subject content
Boys'/girls' subject	Relevancy for future career

Table 7.34 Male and female teachers' views of pupils' reasons for choosing science, clustered by elementary linkage analysis

Men (N=111)	Women (N=53)
1. Relevancy for future career	1. Parental influence
Parental influence	Tradition
Tradition	Boys'/girls' subject
Boys'/girls' subject	
	2. Relevancy for future family life
2. Relevancy for future family life	Like teaching style
Like subject content	Expect to like teaching group
Like teaching style	Like teacher
	Find subject easy
3. Expect to like teaching group	
Like teacher	3. Like subject content
Find subject easy	Relevancy for future career

in general, rather than the separate subjects.

The ten reasons commonly used by teachers to explain pupil uptake of science subjects fall into only two clusters when applied to both boys' and girls' reasons for choosing science (Table 7.33). One cluster refers to subject and teacher characteristics, the other to out-of-school influences. However, the composition of the latter cluster for boys and for girls is not identical. The teachers linked relevance of science for a future career with the social influences when considering boys' choice of science, but relevance for future family life was linked with the same social influences when the reasons why girls choose science was considered.

Differences between the dimensions underlying male and female teachers' replies were also investigated. The replies of both groups of teachers fell into three clusters, but their composition was different for the two sexes (Table 7.34). Male teachers linked career relevance with social influences such as tradition and parental influence, whilst female teachers linked career relevance with the knowledge acquired as a result of studying science.

7.3.3.5 Summary

1. Science teachers believe that most pupils' choice of science results from a mixture of school-based factors (e.g. liking of the subject content, liking of the teacher), and out-of-school pressures (e.g. the subject's relevancy for future careers, parental influence).

2. Science teachers tend to believe that girls are influenced more by social and career considerations when choosing biology than are boys.

3. Tradition and the masculine connotations of science are believed to influence boys more frequently than girls in their choice of science. Besides the simple influence of pupil sex upon these two variables, analysis of variance revealed that pupil sex also interacts with the science subject under consideration to produce more complex effects.

Table 7.35 Ranking and mean rating of factors which influence pupils to drop science (N=164)

Rank	Variable	$\bar{X}$	SD
1	Find subject difficult	3.78	0.72
2	Dislike subject content	3.63	0.67
3	Dislike teaching style	3.17	0.67
4	Irrelevancy for future career	3.16	0.89
5	Dislike teacher	3.04	0.73
6	Irrelevancy for future family life	2.68	1.03
7=	Expect to dislike teaching group	2.67	0.85
7=	Parental influence	2.67	0.77
9	Tradition	2.44	0.99
10	Boys'/girls' subject	2.28	1.11

Most strikingly, science teachers perceive that boys often choose physics because it is a boys' subject and because boys have traditionally chosen to study physics.

4. Teachers believe that pupils choose different subjects for different reasons. They indicated that pupils are more likely to choose physics and chemistry than biology because of career considerations, parental influence and tradition. Biology is more often chosen for its relevance to family life and because pupils like the subject's contents.

5. Cluster analysis suggests that teachers link career relevancy with social influences when considering boys' choice of science, but they substitute relevancy for family life when considering girls' choice of science.

7.3.4 Reasons for Dropping Science

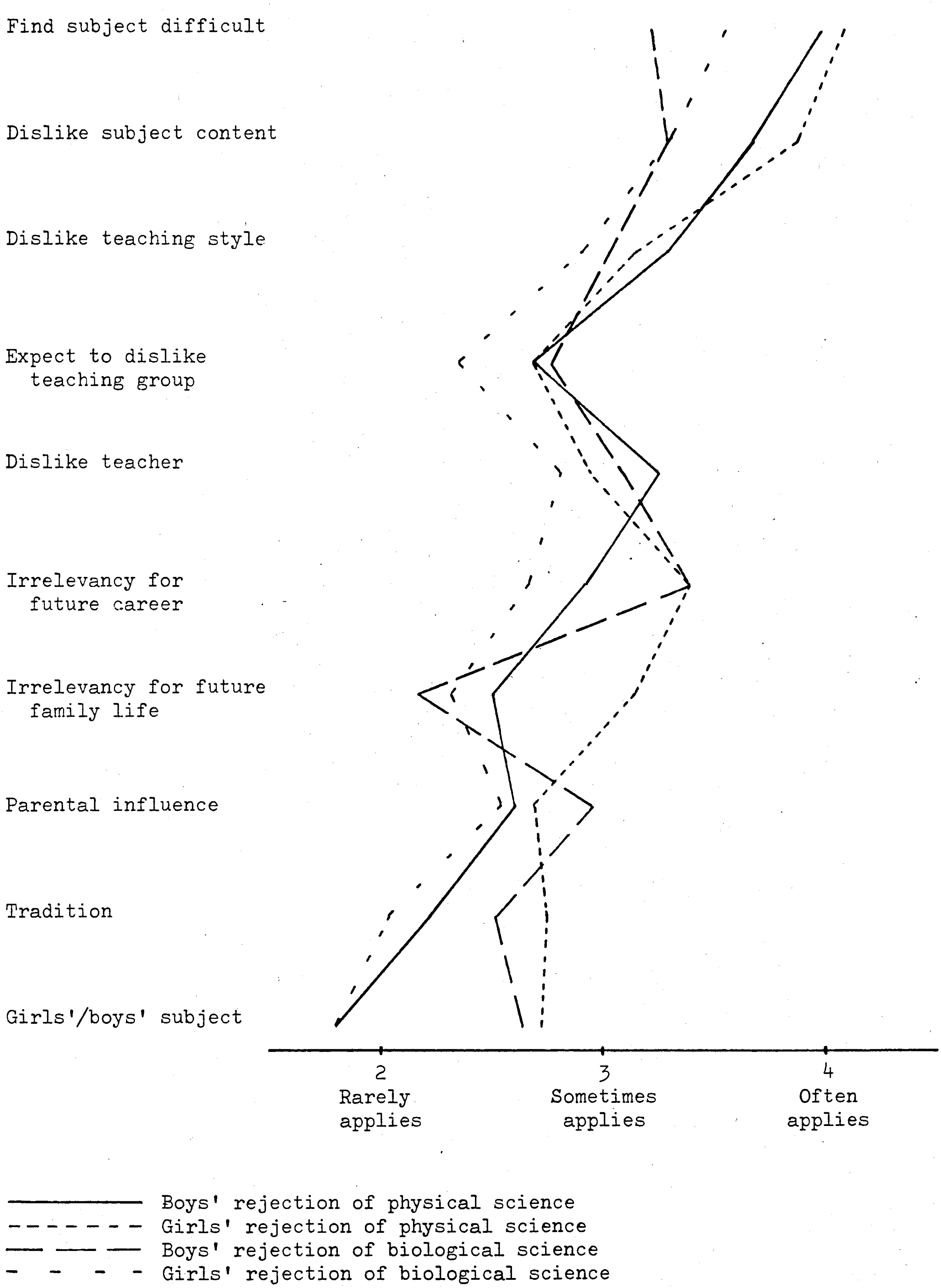
Verbal replies were converted to numerical data as described in section 7.3.3. Thus the higher the number assigned to a variable, the more frequently that variable was thought to influence subject choice adversely.

7.3.4.1 Relative importance of variables

The replies for each reason were averaged and then the mean ratings were ranked in order of frequency to show clearly those reasons that the teachers considered were most important in causing pupils to drop science, and those reasons believed to influence the rejection of science least (see Table 7.35).

Table 7.35 shows that teachers believe that the most important factors influencing pupils to drop science concern, either directly or indirectly, the nature and associated characteristics of science itself. The judgement by pupils that science is irrelevant for their future career is believed to be the most frequently operating reason that is not closely associated with school-based factors. Teachers do not consider

Figure 7.7 Pupils' reasons for dropping science subjects as viewed by science teachers



that diffuse social factors, such as tradition and sex typing of a school subject, are very important influences upon pupils' decisions to drop science subjects.

7.3.4.2 Effect of pupil sex

As a preliminary stage in looking for differences between the motivation patterns believed to operate when boys and girls drop science, the mean ratings associated with boys and girls for each reason were plotted for physical science and biological science (Figure 7.7). The graphs show that not all the reasons are believed to operate with equal frequency, and that the reasons are not all equally appropriate for the two science areas. For example, subject difficulty and subject content are thought to turn more pupils from physics and chemistry than from biology.

Figure 7.7 clearly displays the effect of pupil sex within each subject area. The graphs showing the frequency with which the factors influence boys and girls to drop physical science subjects are very similar, with the exception of large, statistically significant differences for irrelevancy for future career, irrelevancy for future family life, tradition and boys'/girls' subject ($p<0.05$, two-tailed t test). All of the variables were judged to be more important in influencing girls' decisions to drop physics and chemistry, with the exception of dislike of the teacher and his or her teaching style. Turning to biology, there are greater differences between the perceived motivation patterns associated with boys and with girls, although only two reasons, irrelevancy for future career and boys'/girls' subject, produced differences which were statistically significant ($p<0.01$). In contrast to the physical sciences, all of the variables were judged to be more important in influencing boys' decisions to drop biology, with the exception of subject difficulty and irrelevancy for future family life.

Table 7.36 Pupils' reasons for dropping science subjects as viewed by science teachers

Reason	Pupil sex	Phys	Chem	Biol	Analysis of variance Significance: Pupil sex	Subject	Interaction
Find subject difficult	Boys	4.19	3.85	3.22	-	xxx	-
	Girls	4.00	4.13	5.54			
Dislike subject content	Boys	3.69	3.65	3.30	-	xxx	-
	Girls	3.92	3.83	3.29			
Dislike teaching style	Boys	3.50	3.19	3.04	-	x	-
	Girls	3.16	3.17	2.89			
Expect to dislike teaching group	Boys	2.81	2.65	2.78	-	-	-
	Girls	2.68	2.73	2.36			
Dislike teacher	Boys	3.25	3.23	3.09	x	-	-
	Girls	2.76	3.13	2.82			
Irrelevancy for Future career	Boys	2.75	3.04	3.39	-	-	xxx
	Girls	3.48	3.33	2.68			
Irrelevancy for future family life	Boys	2.44	2.54	2.17	xx	xxx	-
	Girls	3.40	2.93	2.32			
Parental influence	Boys	2.50	2.65	2.96	-	-	-
	Girls	2.76	2.67	2.54			
Tradition	Boys	2.31	2.35	2.52	-	-	x
	Girls	2.92	2.60	2.04			
Boys'/Girls' subject	Boys	2.06	1.65	2.65	-	-	xxx
	Girls	2.92	2.57	1.79			
N	Boys	16	26	23			
	Girls	25	30	28			

xxx Significant at 0.1% level
xx Significant at 1% level
x Significant at 5% level
- Not significant

7.3.4.3 Interaction between pupil sex and science subject

Interactions between pupil sex and science subject area can also be discerned in Figure 7.7. For example, tradition more often influences boys than girls to drop biology, but the reverse is true in the physical sciences.

Interactions between pupil sex and science subject were investigated further using analysis of variance. A 2x3 analysis (sex of pupil, science subject) was computed for each motivating reason (see Table 7.36). Teacher sex was not included as an independent variable because preliminary analysis had established that male and female teachers did not differ significantly in their judgements regarding the importance of the different factors.

Pupil sex produced two main effects. The teachers believed that boys were more likely to drop science because they did not like the teacher than were girls. However, the fact that science might be viewed as being irrelevant to future family life is considered to turn more girls than boys from science. Science subject under consideration produced four main effects. The teachers believed that pupils are more likely to drop physics and chemistry than biology because they find the subject difficult and they do not like the subject content. In addition, physics is more often dropped, than the other two sciences, because pupils dislike the style of teaching and they consider the subject to be irrelevant to their future family life. Biology is least frequently dropped for the two foregoing reasons. All three interaction effects are due to the teachers' perception that girls more often than boys drop physics and chemistry, whilst boys more often drop biology, because the subject is judged to be irrelevant for their future careers, the subject is associated with the opposite sex, and because traditionally few pupils of their sex have studied the subject.

Table 7.37 Reasons why boys and girls drop science, clustered by elementary linkage analysis

Boys (N=73)	Girls (N=91)
1. Find subject difficult Dislike subject content Dislike teaching style Dislike teacher Expect to dislike teaching group	1. Parental influence Tradition Boys'/girls' subject
2. Parental influence Tradition Boys'/girls' subject	2. Irrelevancy for future career Irrelevancy for future family life
3. Irrelevancy for future career Irrelevancy for future family life	3. Dislike teaching style Dislike teacher Expect to dislike teaching group
	4. Find subject difficult Dislike subject content

7.3.4.4 Cluster analysis

The existence of interrelationships between the factors and underlying dimensions was investigated using elementary linkage analysis. The ten reasons commonly advanced as to why pupils drop science fell into three clusters when boys were considered and four clusters for girls. This indicates that the teachers perceived girls' motivation for dropping science to be more highly differentiated than boys' motivation. Two of the dimensions that emerged for girls referred to features of the subject and the teaching environment, but they were combined together into a single dimension for boys. The other two clusters for both boys and girls referred to social influences and relevancy of science for adult life.

Differences between the dimensions underlying male and female teachers' replies were also investigated. The replies of men fell into the same clusters as those obtained for boys, and the replies of women were similar to those for girls (see Table 7.37). This suggests that female teachers differentiate more between subject and teacher characteristics than do male teachers.

7.3.4.5 Summary

1. Science teachers believe that the most important factors influencing pupils to drop science concern, either directly or indirectly, the nature and associated characteristics of science itself. Subject difficulty and dislike of a subject's content were judged to be the two most frequently operating deterrents.

2. Teachers believe that pupils drop different subjects for different reasons. They indicated that pupils are more likely to drop physics and chemistry than biology because of subject difficulty and dislike of subject content. In addition, physics is most likely to be dropped because pupils dislike the style of teaching and they consider the subject to be irrelevant to their future family life.

3. According to science teachers, boys are more likely than girls to drop science because they do not like the teacher, and girls are more likely to drop science because of its irrelevancy to their future family life.

4. Analysis of variance produced three interaction effects. Briefly, teachers perceive that girls more often than boys drop physics and chemistry, whilst boys more often drop biology, because the subject is judged to be irrelevant for their future career, the subject is associated with the opposite sex, and because traditionally few pupils of their sex have studied the subject.

7.3.5 Conclusions

1(a) Science teachers believe that internal factors, such as a pupil's effort, ability and interest in the subject, contribute more to success in science than do external factors, such as family support, out-of-class experience (Table 7.25). Similar attribution patterns are used to explain the success of both boys and girls in each of the three main science subject, with one exception. Pupil effort is judged to contribute significantly more to girls' success in physics than to boys' success. Taken altogether, these results signify that Hypotheses Ten and Eleven should be rejected, in spite of some support for Hypothesis Eleven. Science teachers are equally likely to attribute boys' and girls' success at science to stable internal factors, e.g. ability. Science teachers are equally likely to attribute girls' and boys' success at science to unstable factors, e.g. effort or luck.

(b) Science teachers believe that internal factors, such as lack of effort, poor attitude and lack of interest in the subject, contribute most to failure in science (Table 7.29). Similar attribution patterns are used to explain the failure of both boys and girls in each of the three main science subjects, except that subject difficulty is believed to contribute more to girls' failure in physics, and lack of relevant out-of-class experience to boys' failure in biology. These results indicate that Hypotheses Twelve and Thirteen should be rejected. Science teachers are equally likely to attribute boys' and girls' failure at science to unstable factors, e.g. lack of effort or bad luck. Science teachers are equally likely to attribute girls' and boys' failure at science to stable internal factors, e.g. lack of ability.

2. (Relates to Question Fifteen, Do science teachers believe that boys and girls are differently motivated when they choose either to continue studying science or to drop science at 14+?)

(a) Science teachers believe that boys and girls tend to choose science subjects for very similar reasons. A pupil's choice of science is thought to result from a mixture of school-based factors (e.g. liking of the subject content, liking of the teacher), and out-of-school pressures (e.g. the subject's relevancy for future careers, parental influence) (Table 7.31). However, tradition and the masculine connotations of science are believed to influence boys more frequently than girls in their choice of science (Table 7.32). These two variables are thought not to be equally influential in persuading boys to choose the different science subject. Science teachers believe that they most frequently operate to determine boys' choice of physics.

(b) Science teachers believe that the most important reasons causing pupils to drop science are the same for both girls and boys. Pupils most frequently drop a science subject because they find it difficult or because they dislike it (Table 7.35). Sex differences are thought to occur in the operation of some of the less important causes. Teachers think that boys are more likely than girls to drop science because they do not like the teacher, and girls are more likely to drop science because of its irrelevancy to their future family life. The influence of some factors is determined by both a pupil's sex and the science subject under consideration. Thus teachers believe that girls more often than boys drop physics and chemistry, whilst boys more often drop biology, because the subject is judged to be irrelevant for their future career, the subject is associated with the opposite sex, and because traditionally few pupils of their sex have studies the subject (Table 7.36).

7.4 TEACHER EXPECTATION AND TEACHER JUDGEMENT

7.4.1 Marking Exercise

7.4.1.1 Basic considerations

The final form of the marking exercise was completed by 339 science teachers (sample P). Of these, 33 had inappropriate teaching experience for marking chemistry, having never taught chemistry nor integrated science. Consequently their replies were analysed separately, and the findings are reported in section 7.4.1.6. The results reported in sections 7.4.1.2 to 7.4.1.5 are based upon the replies of 306 science teachers. Details regarding the teaching experience of these teachers (reported in Appendix 7.8) indicates that they were well qualified for the task asked of them.

The experimental design assumed that each work sample pair would be attributed to a boy and to a girl an equal number of times, and that every sample pair would be seen in each position in the booklet an equal number of times (see section 4.3.2). These assumptions were largely met, as is shown by Tables A4.13/1 and A4.13/2 in Appendix 4.13.

Since ratings were obtained for each work sample pair when it appeared in all three positions in the booklet, the data can be used to investigate the effect of position upon marks awarded. Findings regarding the influence of the quality of the preceeding work upon the ratings given to subsequent work samples is reported in Appendix 7.9.

The effect of sample, pupil and teacher characteristics upon the marks awarded to the work samples forms the basis of the majority of the analyses that were performed. The influence of the standard of the work samples was not investigated extensively, although mention is made in

Table 7.38 Mean grades awarded to 'boy' and 'girl' on each variable for each sample pair by appropriately experienced science teachers

Variable	Pupil sex	Sample pair: Good	Average	Poor
Standard	Boy	2.74	1.83	1.36
	Girl	2.66	1.82	1.29
Mark merited	Boy	7.59	5.36	3.71
	Girl	7.43	5.05*	3.73
Mark given	Boy	7.73	5.94	4.71
	Girl	7.73	5.56**	4.77
Neatness	Boy	4.35	2.82	3.22
	Girl	4.35	2.71	3.31
Effort involved	Boy	4.21	3.27	3.38
	Girl	4.19	3.35	3.44
Grammar & spelling	Boy	3.79	3.21	3.23
	Girl	3.84	3.04	3.23
Scientific accuracy	Boy	3.80	2.66	1.67
	Girl	3.61*	2.38**	1.60
Understanding of principles	Boy	4.02	2.77	1.51
	Girl	3.71**	2.49*	1.55
Clarity of explanation	Boy	3.30	3.09	2.57
	Girl	3.24	3.02	2.61
Standard of diagram	Boy	4.53	2.78	1.83
	Girl	4.52	2.62*	1.77
Aptitude for science	Boy	4.27	3.02	2.21
	Girl	4.09*	2.81*	2.01*
Attitude towards science	Boy	3.93	3.78	3.64
	Girl	3.62***	3.41***	3.44*
Interest in science	Boy	3.92	3.91	3.47
	Girl	3.56***	3.34***	3.08***
O level suitability	Boy	3.92	2.57	2.06
	Girl	3.59**	2.37*	1.84*
CSE suitability	Boy	3.61	3.99	3.51
	Girl	3.74	3.60***	3.17**
N	Boy	136	183	146
	Girl	170	126	160

* Significant at 5% level
** Significant at 1% level
*** Significant at 0.1% level

Appendices 7.10 and 7.11. The effect of teacher and pupil characteristics, both singly and in combination, upon the marks awarded to the work samples, constitutes the bulk of the results reported in sections 7.4.1.2 to 7.4.1.6. Analyses focus upon the importance of a person's sex to the rating procedure.

Besides investigating the influence of teacher and pupil sex upon the ratings awarded to individual variables, their influence upon the broad dimensions underlying the variables is also investigated. However, initial analysis had first to establish the nature of the dimensions. This work is reported in Appendix 7.12.

7.4.1.2 Effect of pupil sex upon marks awarded

Mean grades awarded to 'boys' and 'girls' for each item on each sample pair are recorded in Table 7.38. It can be seen that, regardless of the standard of the work, the teachers generally rated the work and personal attributes of boys more highly than those of girls. The only work characteristic on which girls tended to be favoured was the amount of effort which had been expended in producing the work. In all, 33 of the 45 comparisons recorded in the table show that work attributed to a boy received higher mean ratings than the same work attributed to a girl. If pupil sex was an inconsequential variable, one would expect boy's work to receive higher ratings in half of the comparisons. To determine whether the observed frequency differs significantly from this expected frequency, χ^2 was calculated (see Appendix 7.13). The value obtained was 9.8, which is significant at the 1% level. Hence it is unlikely that the marks were awarded independently of pupil sex.

Having hypothesized that higher marks would be awarded to the work of a boy than to the work of a girl, a one-tailed t test was used to determine the significance of the difference between the boy's means and the girl's means. 21 of the 45 comparisons were significantly different at the 5% level. This can be compared to the ratio of 2.25/60 which

Table 7.39 Effect Sizes (d) of grades awarded to a boy compared to grades awarded to a girl, by standard of work

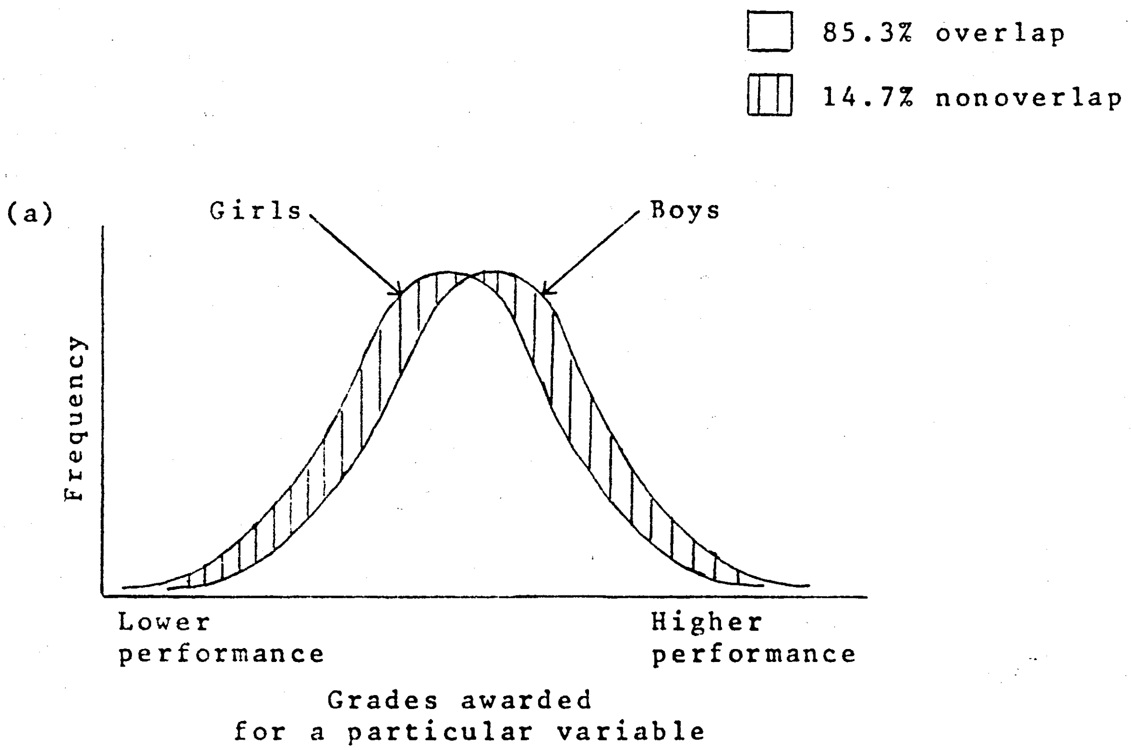
Variable	Standard of Work: Good	Average	Poor	Mean d	U_3 % of girls whose grades were exceeded by 50% boys
Standard	0.15	0.02	0.13	0.1	54.0
Mark merited	0.11	0.20	-0.01	0.1	54.0
Mark given	0	0.29	-0.04	0.1	54.0
Neatness	0	0.15	-0.10	0.0	50.0
Effort involved	0.03	-0.10	-0.08	0.1	54.0
Grammar and spelling	-0.06	0.18	0	0.0	50.0
Scientific accuracy	0.21	0.30	0.10	0.2	57.9
Understanding of principles	0.31	0.26	-0.06	0.2	57.9
Clarity of explanation	0.06	0.08	-0.04	0.0	50.0
Standard of diagram	0.01	0.19	0.08	0.1	54.0
Aptitude for science	0.22	0.27	0.26	0.3	61.8
Attitude towards science	0.39	0.46	0.23	0.4	65.5
Interest in science	0.45	0.66	0.45	0.5	69.1
O level suitability	0.34	0.21	0.26	0.3	61.8
CSE suitability	-0.10	0.41	0.34	0.2	57.9

would be expected to occur by chance. On four of the variables - aptitude for science, attitude towards science, interest in science and O level suitability - boys were rated significantly higher than girls across all three sample pairs.

A statistically significant result implies that some phenomenon exists, but it gives no indication of the magnitude of the phenomenon. To assess the educational significance, as opposed to the statistical significance of the above results, the Effect Size (ES) index 'd' can be calculated. (See Appendix 7.6 for a fuller discussion of ES indices)

The d values associated with the differences between the mean grades received by boys' work and girls' work are presented in Table 7.39. In evaluating d values, Cohen (1977) suggests that d = 0.2 be considered the threshold of significance, indicating that the independent variable accounted for 1% of the variance in the dependent variable. Thus d = 0.2 can be considered a small effect size, d = 0.5 can be described as a moderate effect size, and Cohen suggests that d = 0.8 should be regarded as a large effect size. To aid the interpretation of d values, Cohen provides tables which convert d values to percentage overlap between the two groups. A d value of 0.2 indicates that 14.7% of the combined area covered by both distributions (boys and girls) is not overlapping (see Figure 7.8a). A d value of 0.3 gives a corresponding percentage non-overlap of 21.3, and a d value of 0.5 gives 33.0% nonoverlap. (A complete table is reproduced in Appendix 7.6) Using an alternative measure of overlap, a d value of 0.2 signifies that 57.9% of the girls received lower ratings than the top 50% of boys (see Figure 7.8b). Corresponding percentages for d values of 0.3 and 0.5 are 61.8 and 69.1 respectively. The d value of 0.5 obviously indicates a fairly strong degree of association between the perceived interest shown by a pupil and his or her sex. Although the smaller d values of 0.2 and 0.3 signify considerable overlap between the ratings given to boys and girls, there is a discernible displacement between the two distributions. It should

Figure 7.8 Percentage overlap between grades awarded to boys and girls when d = 0.2



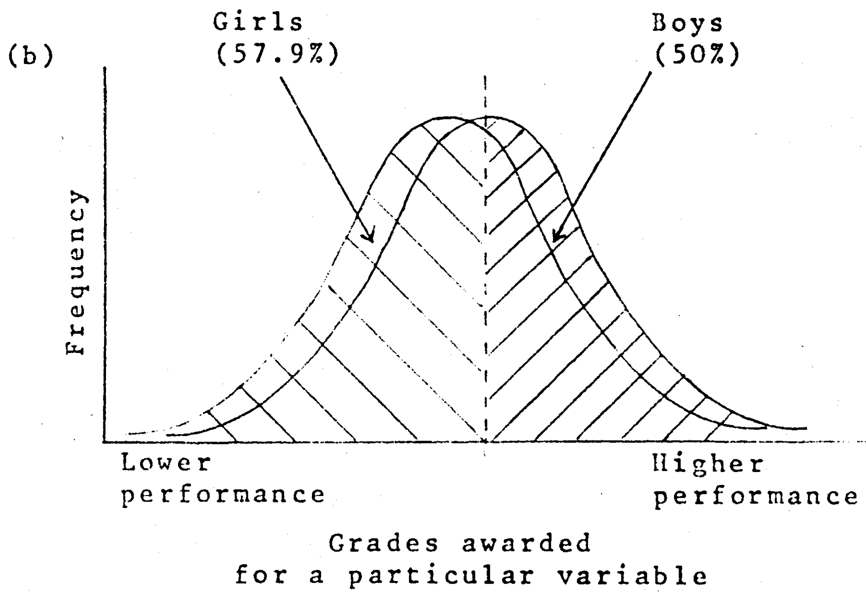


Table 7.40 Mean grades awarded by male and female teachers to each sample pair for each variable

Variable	Teacher sex	Sample pair: Good $\bar{x}$	Average $\bar{x}$	Poor $\bar{x}$
Standard	Male	2.68	1.79	1.34
	Female	2.73	1.90	1.29
Mark merited	Male	7.41	5.09	3.73
	Female	7.68	5.53	3.69
Mark given	Male	7.67	5.72	4.70
	Female	7.88	5.96	4.84
Neatness	Male	4.31	2.79	3.27
	Female	4.40	2.76	3.28
Effort involved	Male	4.14	3.30	3.43
	Female	4.31	3.32	3.40
Grammar and spelling	Male	3.82	3.08	3.20
	Female	3.82	3.26	3.30
Scientific accuracy	Male	3.63	2.53	1.65
	Female	3.83	2.58	1.59
Understanding of principles	Male	3.72	2.56	1.57
	Female	4.11	2.83	1.46
Clarity of explanation	Male	3.24	2.98	2.58
	Female	3.35	3.23	2.62
Standard of diagram	Male	4.51	2.76	1.82
	Female	4.55	2.63	1.75
Aptitude for science	Male	4.11	2.89	2.17
	Female	4.29	3.05	1.98
Attitude towards science	Male	3.80	3.59	3.49
	Female	3.68	3.72	3.64
Interest in science	Male	3.72	3.62	3.22
	Female	3.71	3.80	3.39
O level suitability	Male	3.74	2.44	1.95
	Female	3.72	2.60	1.94
CSE suitability	Male	3.67	3.79	3.25
	Female	3.68	3.89	3.46
N	Male	202		
	Female	101		

not be forgotten that theoretically the two distributions would be expected to overlap exactly.

7.4.1.3 Effect of teacher sex upon marks awarded

Mean grades awarded by male and female teachers to each sample pair for each variable are recorded in Table 7.40. It can be seen that women gave higher mean ratings than men in 31 of the 45 comparisons appearing in the table. This observed frequency can be compared with an expected frequency of 22.5/45, assuming no relationship between a teacher's sex and the marks that teacher awards. The resulting χ^2 value of 6.42 is significant at the 5% level. This suggests that a relationship probably does exist between marks awarded and a teacher's sex.

The significance of the difference between the female teachers' mean ratings and the male teachers' mean ratings was not investigated using t tests, as the data had already been analysed using t tests when comparing the marks received by male and female pupils. The use of multiple t tests is not a satisfactory way of analysing data for maximum clarity of results. Moreover, the comparisons tested by the t tests could not be regarded as completely independent as there would be a tendency for the tests to refer to redundant, overlapping aspects of the data (Hays, 1974). This could lead to difficulties in interpreting the results. Analysis of variance is considered to be the appropriate statistical technique for comparing several sets of categories simultaneously, and also for detecting interactions between the categories.

7.4.1.4 Analysis of variance

Before proceeding to investigate, by analysis of variance, the simultaneous effect of pupil sex and teacher sex on the ratings given to each sample pair, it was first necessary to gain some idea of whether teacher sex or pupil sex also interacted with the standard of the work being marked. If such was the case, then the differences in mean scores

Table 7.41 Analysis of variance of factor scores for work of different standards

Factor	Analysis of variance	Teacher sex	Standard of Work: Good Boy	Good Girl	Average Boy	Average Girl	Poor Boy	Poor Girl
Scientific content		Male	18.26	18.26	14.32	12.95	10.33	10.67
		Female	19.91	18.56	14.78	14.30	10.78	10.31
	Pupil sex		-		xx		-	
	Teacher sex		-		-		-	
	Interaction		-		-		-	
Potential		Male	11.60	11.44	9.61	8.41	7.59	7.12
		Female	12.11	11.37	9.59	9.54	8.03	6.86
	Pupil sex		-		xxx		xx	
	Teacher sex		-		-		-	
	Interaction		-		x		-	
Affective response		Male	7.82	7.29	7.60	6.65	7.02	6.41
		Female	7.87	7.00	7.89	6.92	7.37	6.80
	Pupil sex		xxx		xxx		xx	
	Teacher sex		-		-		-	
	Interaction		-		-		-	
Presentation		Male	16.67	16.87	12.12	11.66	11.76	11.69
		Female	17.22	16.96	12.06	11.82	11.59	11.82
	Pupil sex		-		-		-	
	Teacher sex		-		-		-	
	Interaction		-		-		-	
N		Male	89	113	119	83	103	99
		Female	46	55	63	38	41	60

x Significant at 5% level

xx Significant at 1% level

xxx Significant at 0.1% level

\- Not significant

for one independent variable, i.e. teacher sex or pupil sex, would be larger at some rather than other levels of the second independent variable, i.e. standard of work. Analysis of variance is an effective method of detecting such interactions between independent variables. Unfortunately the inclusion of the variable 'standard of work' in analysis of variance computations was impeded by the design of the experiment and the organisation of the data. However, there was the possibility of including it if a reduced sample size was employed. To access the feasibility of choosing this option, graphs were drawn to display the existence and nature of any first order interactions between standard and pupil sex, and between standard and teacher sex (see Appendix 7.11 for full details).

Most of the curves in most of the graphs were broadly parallel. This implies either that there was no interaction between the independent variables or that only very weak interactions occurred. Therefore it was judged that so long as the main effect of standard of work upon marks awarded was not overlooked, then the omission of sample standard from subsequent ANOVA (analysis of variance) analyses would not seriously detract from any resulting interactions. Consequently, the ratings awarded to each sample pair for each variable were subjected to 2x2 (sex of pupil, sex of teacher) analysis of variance. Further details regarding the form of analysis of variance used can be found in Appendix 7.7.

The results of the analysis (recorded in full in Appendix 7.14) showed that pupil sex produced a number of main effects, especially for understanding of principles, attitude towards science, interest in science, O level suitability and CSE suitability, with a boy being given higher ratings on these variables than a girl. Teacher sex produced fewer main effects, although they were uniformly in the direction of female teachers having awarded higher grades than male teachers. There were very few significant interactions between pupil sex and teacher sex.

Table 7.42 Teacher sex/pupil sex combinations which gave highest marks

		Pupil sex: Boy	Pupil sex: Girl
Teacher sex	Male	6	2
	Female	30	7

Table 7.43 Teacher sex/pupil sex combinations which gave lowest marks

		Pupil sex: Boy	Pupil sex: Girl
Teacher sex	Male	11	19
	Female	3	12

Where they did occur, examination of the means revealed that it was generally an additive effect, with high marks being received by boys who had been marked by female teachers, and low marks being given to girls who had been marked by male teachers.

To summarize effectively the above results, a further 2x2 (sex of pupil, sex of teacher) analysis of variance was computed using factor scores. The common factors which underlie teachers' marking practices are described in Appendix 7.12. The results of the analysis (Table 7.41) show that boys were awarded significantly higher mean scores for affective variables across all three standards of work. In addition, boys who produced average and poor standard work were judged to display significantly greater potential for science studies than were comparable girls. Another main effect for pupil sex was found for the sample pair of average standard. The scientific content of work produced by a boy was marked significantly higher than that produced by a girl. No main effects were found for teacher sex, and the only interaction effect referred to the potential for science score associated with the average sample pair. As was found with the individual variables, it was an additive effect with a boy marked by female teachers getting a high mean score, and a girl marked by male teachers getting a low mean score.

7.4.1.5 Combined effect of teacher sex and pupil sex

The analysis of variance results reported in the previous section suggest that the combined effects of teacher sex and pupil sex act in an additive manner, rather than interact. To investigate this effect further, the mean grades recorded in Appendix 7.14 were subjected to an additional analysis. The teacher sex/pupil sex combination which produced the highest mark on each variable for each sample pair was recorded (Table 7.42). Similar counts were made of the combinations producing the lowest marks (Table 7.43). This method of analysis allowed the results of all three sample pairs to be considered simultaneously.

Table 7.44 Mean grades awarded to 'boy' and 'girl' on each variable for each sample pair by inappropriately experienced science teachers

Variable	Pupil sex	Standard of work Good $\bar{x}$	Average $\bar{x}$	Poor $\bar{x}$
Standard	Boy	2.00	1.76	1.25
	Girl	2.74*	1.67	1.06
Mark merited	Boy	5.20	5.01	3.75
	Girl	7.74***	5.00	2.88
Mark given	Boy	5.60	5.41	4.44
	Girl	7.87***	5.60	3.75
Neatness	Boy	4.00	2.65	3.25
	Girl	4.65***	2.56	2.76
Effort involved	Boy	3.60	3.24	3.44
	Girl	4.30**	3.25	3.18
Grammar and spelling	Boy	2.60	3.24	3.19
	Girl	3.74***	2.94	2.94
Scientific accuracy	Boy	2.70	2.47	1.69
	Girl	4.00***	2.88	1.70
Understanding of principles	Boy	3.10	2.59	1.56
	Girl	4.22*	2.56	1.44
Clarity of explanation	Boy	2.20	2.59	2.44
	Girl	3.39*	2.81	2.25
Standard of diagram	Boy	4.40	2.82	2.38
	Girl	4.65	2.69	1.63**
Aptitude for science	Boy	3.20	3.00	2.06
	Girl	4.30*	2.73	1.81
Attitude towards science	Boy	3.60	3.59	3.25
	Girl	3.78	3.69	3.24
Interest in science	Boy	3.40	3.82	3.13
	Girl	3.78	3.69	2.88
O level suitability	Boy	3.30	2.53	2.13
	Girl	4.00*	2.50	1.88
CSE suitability	Boy	3.80	3.88	3.44
	Girl	3.74	3.81	3.18
N	Boy	10	17	16
	Girl	23	16	17

* Significant at 5% level
** Significant at 1% level
*** Significant at 0.1% level

χ^2 values were calculated to establish whether the frequencies recorded in Tables 7.42 and 7.43 differ significantly from 11.25(45/4), the mean expected frequency assuming that marks are awarded independently of teacher sex and pupil sex. A value of 42.91 was obtained for Table 7.42, and Table 7.43 gave 11.44. Both these values indicate that the null hypothesis can be rejected at the 0.1% level. It is very probabl that teacher sex and pupil sex act together to determine marks awarded.

7.4.1.6 Effect of teaching subject upon marks awarded

The results presented in section 7.4.1.2 clearly show that written work attributed to a boy was often given significantly higher grades than identical work attributed to a girl when marked by those science teachers in the sample who were familiar with the topic that they were marking. This sex biased marking pattern occurred across all three sample pairs which indicates that it operated regardless of the standard of the work being marked. Since a small proportion of the sample of teachers had no experience of teaching chemistry or integrated science, the opportunity existed of investigating and comparing the marking patterns of these teachers who had inappropriate teaching experience for the task asked of them.

Table 7.44 records the mean grades awarded to 'boys' and 'girls' for each item on each sample pair by the science teachers with inappropriate teaching experience. When the standard of work was high, these teachers favoured the work of a girl over that of a boy on all the variables marked, with the exception of CSE suitability. However, this variable had been found to be rather ambiguous when applied to the good work sample. Many teachers indicated that the pupil was unsuited for CSE courses, believing that the pupil was better suited for O level courses. Thus a low rating on CSE suitability could well be interpreted as a favourable assessment of the pupil's potential for science studies.

Statistical analyses were computed to determine the significance and magnitude of the tendency to award higher marks to the work of a girl when the standard of work was high. χ^2, with Yates correction (see Appendix 7.13), was calculated to determine whether the observed frequency of 14/15 comparisons favouring the girl's work differs significantly from the expected frequency of 7.5/15. The value obtained was 9.6, which is significant at the 1% level. This χ^2 value indicates that it is unlikely that grades were awarded independently of pupil sex. This conclusion is further supported by t values. Two-tailed t tests were used to determine the significance of the difference between the girl's means and the boy's means. Eleven of the 15 comparisons were significantly different at the 5% level or better. This ratio can be compared to the ratio of 0.75/15 which would be expected to occur by chance.

For the sample pair of average standard, higher grades were awarded to the work of a girl in only five of the comparisons, and the work of a boy received higher grades in the remaining ten comparisons. The χ^2 value testing the difference of this ratio from that predicted by the null hypothesis was 1.07, which is not significant. Therefore the null hypothesis that the boy will receive higher ratings in half of the comparisons cannot be rejected. Furthermore, none of the differences between the girl's mean ratings and the boy's mean ratings were significant using a two-tailed t test. These results indicate that the teachers with inappropriate teaching experience did not mark the work of a boy and girl differently when the work was of average standard.

The marking pattern for the sample pair of poor standard differed from that encountered with both the good and average standard work. The inappropriately experienced science teachers awarded the work of a boy higher mean grades for 14 of the 15 items considered. This ratio gave a χ^2 value of 9.6, which is significant at the 1% level. To determine the significance of the differences between the boy's means and the girl's means t values were calculated. Only one t value was significant at the

Table 7.45 Effect sizes (d) of grades awarded to a boy compared to grades awarded to a girl by inappropriately experienced science teachers

Variable	Standard of work Good	Average	Poor
Standard	-1.11	0.15	0.51
Mark merited	-1.31	0.01	0.58
Mark given	-1.30	-0.12	0.44
Neatness	-1.15	0.14	0.51
Effort involved	-1.03	-0.01	0.34
Grammar and spelling	-1.27	0.33	0.32
Scientific accuracy	-1.23	-0.41	-0.01
Understanding of principles	-1.06	0.03	0.19
Clarity of explanation	-1.17	-0.25	0.21
Standard of diagram	-0.38	0.16	0.89
Aptitude for science	-1.12	0.38	0.35
Attitude towards science	-0.21	-0.13	0.01
Interest in science	-0.44	0.16	0.28
O level suitability	-0.78	0.04	0.29
CSE suitability	0.05	0.08	0.29

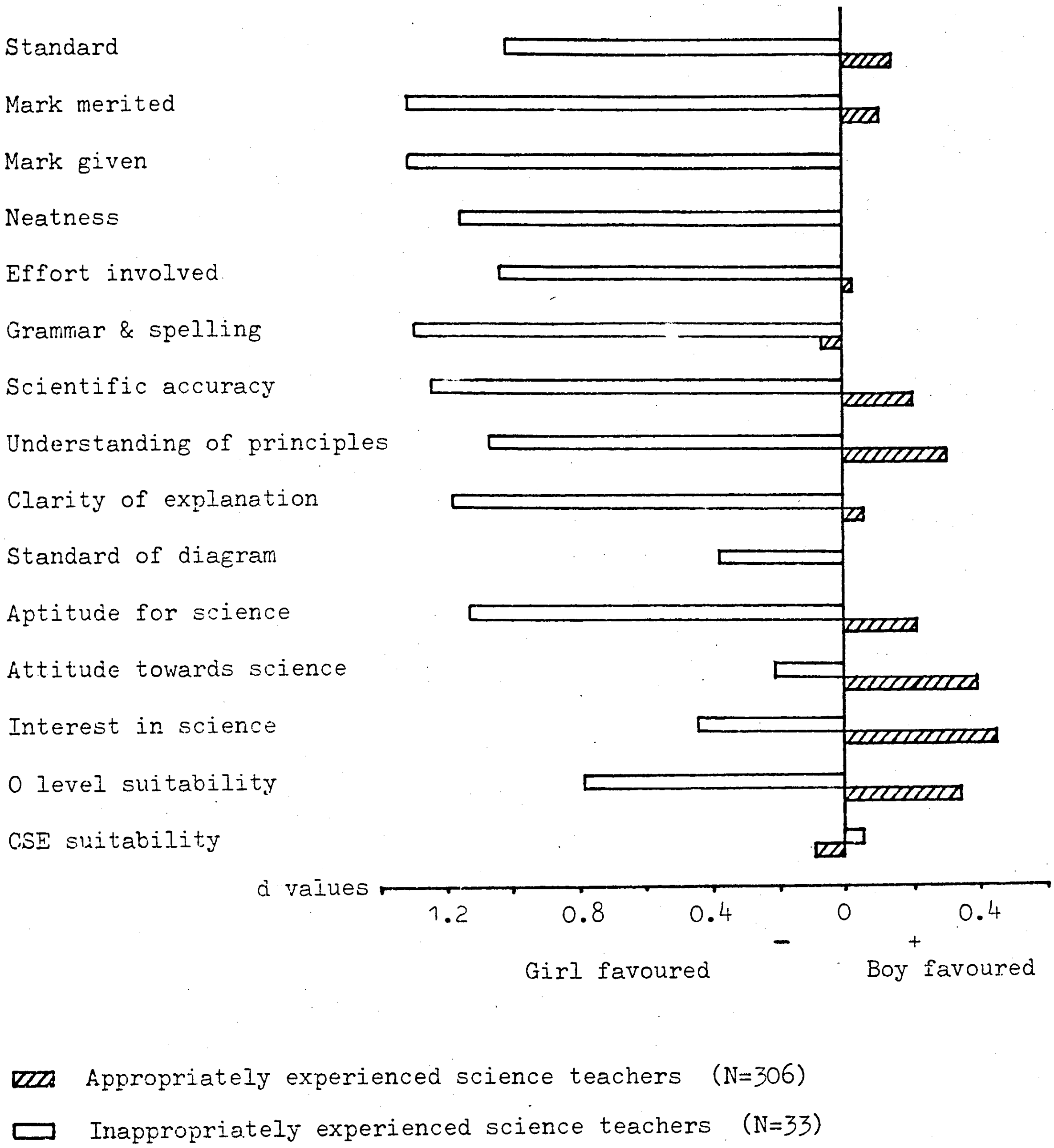
5% level. However, the small size of the sample (N=33) should not be overlooked. With a larger sample size, several of the other differences could be expected to reach statistical significance.

Table 7.45 presents the ES (d) values associated with the mean grades recorded in Table 7.44. These d values indicate the magnitude of the difference between the mean grades awarded to the work of a boy and a girl for each variable. The values clearly emphasize the pattern of biased marking described in the paragraphs above. This biased pattern is most effectively summarized by the mean d values associated with the three sample pairs. These values are -0.90, 0.04 and 0.35 for the good, average and poor sample pairs respectively.

ES values allow direct comparisons to be made with equivalent values obtained from other samples. Thus the d values associated with the marks awarded by the inappropriately experienced science teachers and the appropriately experienced science teachers can be directly compared. Both sets of values for the good sample pair are shown in Figure 7.9. The differences are striking. The inappropriately experienced science teachers rated the work of a girl higher than that of a boy, whilst the reverse sex bias was displayed by the appropriately experienced science teachers. Also, the bias in the marking of the inappropriately experienced science teachers was greater than that displayed by the appropriately experienced science teachers. Finally, the inappropriately experienced science teachers tended to show less sex bias in their marking of those variables on which the appropriately experienced science teachers showed most sex bias. The d values associated with the two affective variables, attitude and interest, illustrate this last point most clearly.

A comparison of the two sets of d values associated with the poor sample pair indicates that both samples of teachers favoured the work of a boy over that of a girl. Furthermore, the inappropriately experienced science teachers displayed greater sex bias in their marking than did

Figure 7.9 Sex bias, illustrated by d values, shown by two groups of teachers when marking work of a high standard



the appropriately experienced science teachers. This last finding is similar to that obtained for the good sample pair, except for the direction of the sex bias.

7.4.1.7 Summary

1. Work attributed to a boy was generally rated higher for scientific accuracy and understanding of principles than identical work attributed to a girl. Furthermore, boys were judged by appropriately experienced science teachers to have significantly more aptitude for science, more favourable attitudes towards science, greater interest in science, and to be more suitable for O level physical science courses.

2. To clarify the effects of pupil sex upon the marks awarded to written work, compound factor scores were subjected to analysis of variance. The results emphasized that the assessment of presentation variables had been totally unbiased. In striking contrast, boys had been awarded significantly higher mean scores for affective variables across all three standards of work. In addition, boys who produced average and poor standard work had been judged to display significantly greater potential for science studies than had similar girls. Lastly, the scientific content of work of an average standard was marked significantly higher when it had been produced by a boy.

3. Female teachers generally gave higher ratings than male teachers.

4. Teacher sex and pupil sex acted together in an additive manner to determine marks awarded. The combination of boy's work marked by a female teacher most frequently produced a generous assessment, whereas girl's work marked by a male teacher most frequently produced a severe assessment.

5. Science teachers, with inappropriate teaching experience for the marking exercise, favoured the work of a girl when the standard of the work was high, favoured the work of a boy when the standard of the work was low, and displayed no sex bias when the work was of average standard.

6. The bias detected in the inappropriately experienced science teachers' marking of the good and poor pieces of work was greater than that displayed by the appropriately experienced science teachers. However, in their marking of the good piece of work, the inappropriately experienced science teachers tended to give less sex biased marks to those variables which elicited the most sex biased marks from the appropriately experienced science teachers.

7.4.2 Conclusions

1. In a marking exercise, science teachers evaluated samples of pupils' work. The work samples and their authors were rated on a number of variables. However, pupil sex was varied so that the same piece of work was presented to half of the teachers as being the work of a girl and to the remaining teachers as being the work of a boy.

(a) Work attributed to a boy was generally rated higher for scientific accuracy and understanding of principles than was work attributed to a girl (Table 7.38). This finding confirms Hypothesis Fourteen. For identical written work, science teachers award higher marks to boys than to girls.

(b) Boy authors were judged to be more suitable for O level physical science courses than were girl authors (Table 7.38). This confirms Hypothesis Fifteen. Based on the evidence of written work, science teachers form higher expectations for boys than for girls, as signified by their judgement of pupils' potential for science.

(c) Boy authors were judged to have significantly more aptitude for science than were girls authors (Table 7.38). This finding provides support for Hypothesis Sixteen. Based on the evidence of written work, science teachers are more likely to judge that a boy, than a comparable girl, possesses cognitive ability that is appropriate for the study of science.

(d) Boy authors were judged to have significantly more favourable attitudes towards science and significantly greater interest in science than were girl authors (Table 7.38). Thus Hypothesis Seventeen should also be accepted. Based on the evidence of written work, science teachers are more likely to judge that a boy's attitude towards science and his interest in science are superior to those of a comparable girl.

(e) Female teachers generally gave higher ratings than male teachers (Table 7.40). This finding confirms Hypothesis Eighteen. When marking identical samples of written work, female teachers give higher marks than do male teachers.

(f) (Relates to Question Sixteen, Does teacher sex interact with pupil sex to further complicate the marks awarded to pupils?) Teacher sex and pupil sex acted together in an additive manner to determine marks awarded. The combination of boy's work marked by a female teacher most frequently produced a generous assessment (Table 7.42), whereas girl's work marked by a male teacher most frequently produced a severe assessment (Table 7.43).

2. Question Eighteen "Do science teachers form more sex differentiated expectations for pupils when less information is available?" was not answered satisfactorily. See Appendix 6.15 for further details.

CHAPTER 8

RELATIONSHIPS BETWEEN THE INDEPENDENT AND DEPENDENT VARIABLES

8.0 Contents

CHAPTER 8

TABLES

Page

FIGURES

One of the aims of the investigation was to identify the existence of relationships between the independent variables and the dependent variables, and within the dependent variables. Although the investigation of some of these relationships was informed by a formal hypothesis and so was planned in advance, many of them were examined on a post hoc basis. Findings arising from the latter type of analysis form the basis of this chapter.

8.1 EFFECT OF TEACHER SEX AND PRINCIPAL TEACHING SUBJECT

8.1.1 Introduction

It was anticipated that only two of the independent variables recorded might consistently produce meaningful and identifiable effects within the four broad topics under investigation. Therefore these two variables - a teacher's principal teaching subject and his/her sex - were studied extensively. This section reports the effects of these two variables upon a number of the dependent variables used to measure the sex typing of science by science teachers and the sex stereotypes held by science teachers in the main study.

The findings arising from one of the final scales used to measure sex typing and another scale that measured sex stereotypes are not included in this section. Teaching subject and teacher sex effects upon a teacher's perceptions of school subject characteristics are reported in section 7.1.1. In that section the replies of science teachers are compared with those of non-science teachers, and so it seemed appropriate that comparisons between science teachers should also be recorded in the same section. The omission of detailed reference to the School Subject Characteristics scale in this section does not impoverish the ensuing discussion of the sex typing of science, since no differences were detected between science teachers' replies on the basis of their sex or principal teaching subject. The second scale omitted from this section

Table 8.1 Effect of teacher sex and teaching subject upon Masculinity Index ratings

Subject rated	Adjective scale	Teacher sex	Mean rating from teachers of: Phys	Chem	Biol	Analysis of variance Significance: Teacher sex	Teaching subject	Interaction
Physics	Hard	Male	2.20	2.15	2.32	-	-	-
		Female	2.33	2.43	2.39			
	Tough	Male	2.89	2.83	2.40	-	-	-
		Female	3.00	2.29	2.57			
	Cold	Male	3.89	2.76	2.64	-	***	-
		Female	3.33	2.79	3.00			
	Remote	Male	3.66	3.22	2.88	-	-	-
		Female	3.00	2.57	3.04			
	Total	Male	12.63	10.95	10.24	-	**	-
		Female	11.67	10.07	11.00			
Chemistry	Hard	Male	2.49	2.32	2.68	**	-	-
		Female	2.50	3.21	3.09			
	Tough	Male	3.00	3.12	3.00	-	-	-
		Female	3.00	3.36	3.22			
	Cold	Male	3.83	3.85	3.48	-	-	-
		Female	3.17	4.21	3.83			
	Remote	Male	3.51	3.54	2.88	-	*	-
		Female	3.33	3.93	3.52			
	Total	Male	12.83	12.83	12.04	*	-	-
		Female	12.00	14.71	13.65			
Biology	Hard	Male	4.34	4.10	3.84	-	-	-
		Female	3.17	4.57	3.67			
	Tough	Male	4.43	4.73	4.60	*	-	-
		Female	3.67	4.07	4.33			
	Cold	Male	4.83	4.76	5.64	-	**	-
		Female	5.00	4.57	5.00			
	Remote	Male	4.86	4.63	5.24	-	-	**
		Female	4.00	5.43	4.96			
	Total	Male	18.46	18.22	19.32	-	-	-
		Female	15.83	18.64	17.96			
	N	Male	35	41	25			
		Female	6	14	23			

* Significant at 5% level
** Significant at 1% level
*** Significant at 0.1% level
\- Not significant

is the Females' Social Roles scale. For this scale, teacher sex was an important independent variable. The statistical comparison performed to investigate the effect of teacher sex was planned in advance on the basis of a hypothesis stated in Chapter 3. Since teacher sex constituted a major independent variable and not just an incidental independent variable, the findings are reported in section 7.2.3.2.

Reference is not made in this section to the effects of a teacher's sex and teaching subject upon his/her attribution patterns, expectations and judgements. Principal teaching subject was used as a major independent variable in analyses of the attribution pattern scales, and so is not considered in the present context. Findings regarding the effect of teacher sex upon the replies to the attribution scales are included in section 7.3, along with the effects of teaching subject. Discussion of the marking exercise is excluded from this section because teacher sex constituted a major independent variable. It was mentioned in a hypothesis stated in Chapter 3, and formed the basis of planned statistical comparisons in Chapter 7. Principal teaching subject did not constitute a relevant independent variable for the marking exercise since teachers were only categorized on the basis of their experience of teaching chemistry and/or integrated science.

The sex typing and sex stereotyping scales discussed in this section were all analysed by analysis of variance to determine the individual and combined effects of teacher sex and teaching subject upon teachers' replies. Each individual item of each scale was subjected to a 2x3 analysis (sex of teacher, principal teaching subject). The technique of analysis of variance is discussed more fully in Appendix 7.7.

8.1.2 Sex typing of science

8.1.2.1 Masculinity Index

Teachers' perceptions of the masculinity of the three science subjects, as indicated by masculinity indices, were influenced by the

Table 8.2 Effect of teacher sex and teaching subject upon perceptions of science subject characteristics

A. Physics					Analysis of variance Significance		
		Mean rating from teachers of					
Characteristic	Teacher sex	Phys	Chem	Biol	Teacher sex	Teaching subject	Inter action
Logical	Male	1.31	1.29	1.23	-	-	-
	Female	1.17	1.27	1.44			
Objective	Male	1.54	1.34	1.38	-	-	-
	Female	1.17	1.53	1.56			
Relevant for careers	Male	1.31	1.44	1.73	-	*	-
	Female	1.50	1.60	1.72			
Relevant for family life	Male	2.37	2.61	2.62	-	-	-
	Female	2.17	2.53	2.68			
Mathematical	Male	1.54	1.34	1.31	-	-	-
	Female	1.33	1.40	1.24			
Wordy	Male	2.54	2.66	2.69	-	-	-
	Female	2.33	2.87	2.60			
Concerned with people	Male	2.69	2.98	3.08	-	**	-
	Female	2.50	3.00	3.16			
Concerned with objects	Male	1.77	1.44	1.38	-	*	-
	Female	1.33	1.33	1.40			
Concerned with social issues	Male	2.71	3.20	3.27	-	**	-
	Female	2.83	3.20	3.16			
Unfamiliar	Male	2.49	2.44	2.19	-	-	-
	Female	2.33	2.40	2.24			
Technical	Male	1.66	1.54	1.54	-	-	-
	Female	1.33	1.47	1.32			
Mechanical	Male	1.91	1.78	1.58	-	-	-
	Female	1.83	1.53	1.44			
Masculine	Male	2.57	2.12	2.15	*	*	-
	Female	2.83	2.40	2.56			
Abstract	Male	2.37	2.07	2.08	*	-	-
	Female	2.67	2.13	2.56			
Impersonal	Male	2.31	2.07	1.54	*	**	-
	Female	2.33	2.07	2.20			

two variables - teacher sex and teaching subject (see Table 8.1). However, the two variables were not equally forceful in influencing perceptions of the gender connotations of all three science subjects.

Teachers' perceptions of the masculinity of physics were affected by their principal teaching subject. The overall physics Masculinity Index score of physics teachers was higher than that of chemistry and biology teachers, indicating that they regarded physics as a less masculine subject than did the other science teachers. Physics teachers' ratings on the individual semantic differential scales only differed significantly from those of other science teachers on the cold-warm scale. Physics teachers saw physics as being less cold than did the other science teachers.

Judgements of the masculinity of chemistry were affected by a teacher's sex. Female teachers gave higher overall chemistry Masculinity Index scores than did male teachers. Differences between women's and men's replies to the individual rating scales were most marked on the hard-soft scale. The women judged chemistry to be much less hard than did the men teachers.

The effects of teacher sex and teaching subject upon the biology Masculinity Index ratings were not so general. Two main effects were identified on the individual rating scales. Female teachers saw biology as being a less tender subject than did male teachers, and biology teachers described biology as being a much warmer subject than did the physical science teachers.

8.1.2.2 Characteristics of Science

The science teachers viewed some aspects of science subjects differently according to their sex and principal teaching subject (see Table 8.2). Physics was viewed as being a more person-oriented subject by physics teachers than by chemistry and biology teachers. For instance, physics teachers indicated that physics is more concerned with people and

Table 8.2 contd.

B. Biology Characteristic	Teacher sex	Mean rating from teachers of Phys	Chem	Biol	Analysis of variance Significance Teacher sex	Teaching subject	Inter-action
Logical	Male	2.15	2.28	1.69	-	***	-
	Female	2.33	2.40	2.04			
Objective	Male	2.06	1.90	1.88	-	-	-
	Female	1.83	1.93	1.92			
Relevant for careers	Male	2.18	2.15	2.19	*	-	-
	Female	2.00	2.07	1.80			
Relevant for family life	Male	1.88	2.13	1.73	**	-	-
	Female	1.83	1.60	1.52			
Mathematical	Male	3.24	3.18	2.88	*	**	-
	Female	2.67	3.13	2.72			
Wordy	Male	1.74	1.75	1.73	-	-	-
	Female	1.50	1.73	1.72			
Concerned with people	Male	1.82	2.10	1.77	-	*	-
	Female	1.50	1.87	1.68			
Concerned with objects	Male	2.59	2.70	2.85	-	-	-
	Female	2.50	2.73	2.92			
Concerned with social issues	Male	2.50	2.38	1.96	-	-	*
	Female	1.50	2.00	2.04			
Unfamiliar	Male	2.74	2.95	3.00	-	-	-
	Female	2.83	3.07	3.28			
Technical	Male	2.82	2.63	2.38	-	-	-
	Female	2.50	2.67	2.64			
Mechanical	Male	3.15	3.05	3.04	-	-	-
	Female	2.83	3.07	3.08			
Masculine	Male	3.26	3.13	2.96	-	-	-
	Female	3.17	3.33	3.28			
Abstract	Male	2.71	2.73	3.04	*	-	-
	Female	2.83	3.33	3.08			
Impersonal	Male	3.00	2.70	3.19	*	-	-
	Female	3.17	3.40	3.24			
N	Male	35	41	26			
	Female	6	15	25			

* Significant at 5% level
** Significant at 1% level
*** Significant at 0.1% level
\- Not significant

with social issues than did the other two groups of teachers. These views were further complemented by physics teachers perceiving physics to be a less impersonal subject and less concerned with objects. It is interesting to note that physics and biology teachers generally held the most extreme views of physics, with chemistry teachers holding intermediate views. This was also true for their views about the relevance of physics for careers. Not surprisingly, physics teachers were most convinced of the value of physics for careers. Even though these careers are presumably male dominated, physics teachers viewed physics as being a less masculine subject than did the other two groups of teachers. Comparing the responses of male and female teachers showed that male teachers of all subjects judged physics to be more masculine than did female teachers. Male teachers of all subjects also rated physics as being a more abstract subject. Male biology teachers judged physics to be more impersonal than did female biology teachers, but this sex difference did not appear among the chemistry and physics teachers.

Biology was viewed more similarly by teachers of all three subjects. However, biology teachers viewed biology as being more logical and mathematical, a couple of masculine traits, than did physical science teachers. Biology teachers also indicated that biology is more concerned with people. The lowest ratings on these three variables were given by chemistry teachers. Turning to sex differences, females were more convinced of the relevance of biology for careers and for family life than were males. As for physics, female teachers also saw biology as being less abstract and less impersonal than did male teachers. The only other sex difference was that female teachers judged biology to be more mathematical than did male teachers.

8.1.2.3 Scientist Stereotypes

A physicist was for the most part viewed very similarly by all science teachers, regardless of their principal teaching subject or sex

Table 8.3 Effect of teacher sex and teaching subject upon perceptions of scientists

A. Physicist	Teacher sex	Mean rating from teachers of Phys	Chem	Biol	Analysis of variance Significance Teacher sex	Teaching subject	Inter-action
Male	Male	2.34	2.32	2.04	**	-	*
	Female	3.50	2.13	2.88			
Good at maths	Male	1.66	1.73	1.42	-	-	-
	Female	1.33	1.40	1.68			
Logical	Male	1.91	1.88	1.69	-	-	-
	Female	1.83	1.40	2.00			
Objective	Male	2.20	2.07	1.88	-	-	-
	Female	1.83	2.20	2.12			
Competitive	Male	3.26	3.05	3.00	-	-	-
	Female	3.17	3.33	2.96			
Unsociable	Male	4.69	4.39	4.69	-	-	-
	Female	5.17	4.13	5.04			
Unemotional	Male	4.11	4.12	4.62	-	-	-
	Female	4.50	3.87	4.52			
Not humanitarian	Male	4.63	4.02	4.12	-	*	-
	Female	5.00	4.00	4.36			
B. Biologist							
Male	Male	4.66	4.46	4.12	-	*	-
	Female	4.50	4.67	4.24			
Good at maths	Male	3.97	3.39	3.85	-	-	-
	Female	4.00	3.80	3.60			
Logical	Male	3.43	3.49	2.77	-	*	-
	Female	2.50	3.53	2.84			
Objective	Male	2.91	3.17	2.50	-	-	-
	Female	2.33	2.60	2.72			
Competitive	Male	3.51	3.83	2.96	-	**	-
	Female	3.50	3.73	3.24			
Unsociable	Male	4.89	4.90	5.50	-	*	-
	Female	5.17	5.27	5.44			
Unemotional	Male	4.89	4.90	5.50	-	-	-
	Female	4.33	4.80	4.92			
Not humanitarian	Male	5.06	4.54	5.38	-	-	**
	Female	5.50	5.80	4.92			
N	Male	35	41	26			
	Female	6	15	25			

* Significant at 5% level, ** Significant at 1% level, - Not significant

(see Table 8.3). Only two main effects were identified. Generally, the belief of male teachers that physicists are male was stronger than that of female teachers. However, chemistry teachers appeared not to follow this pattern. The other effect arose because physics teachers thought that physicists are more humanitarian than did other science teachers.

Several significant differences were detected between the stereotype of a biologist held by physical science teachers and by biology teachers. The physical science teachers were less likely to associate a biologist with the male sex. They were also less likely to view a biologist as being logical and competitive. Interestingly, the biology teachers thought that a biologist was more likely to be sociable than did the physical science teachers.

8.1.2.4 Summary

(Points 2(a) and 3(a) refer to Question Four, Do science teachers with different subject specialities vary in their views about the masculinity of science?)

1. Science teachers' perceptions of the masculine image of science subjects, their perceptions of a range of the characteristics possessed by science subjects, and their perceptions of scientists were all influenced by their principal teaching subject and their sex. Overall ratings on the first two scales mentioned were equally influenced by teacher sex and teaching subject, but teaching subject proved to be a more potent variable affecting teachers' perceptions of scientists than did teacher sex. Comparing the ratings awarded to individual science subjects and individual scientists, it was found that the independent variable of teaching subject had greater effect upon perceptions of physics than it did upon perceptions of physicists, whereas the reverse was the case for biology and biologists.

2(a) Physics teachers perceived physics to be a less masculine subject than did chemistry and biology teachers. This was indicated by the less

Table 8.4 Physical science and biology teachers' mean ratings of pupils' preferences for subject characteristics

Characteristic	Girls' preference			Boys' preference	
	Phy. Sc. teachers		Biology teachers	Phy. Sc. teachers	Biology teachers
Practical-theoretical	4.08		3.32	2.53	2.11
Numerical-verbal	5.39		5.16	3.39	3.21
Science-arts	4.97	*	4.26	2.81	2.53
Logical-intuitive	4.56		4.63	3.14	2.89
Masculine-feminine	5.28	*	4.58	2.36	2.37
Factual-opinionative	3.44		3.26	2.83	2.63
Routine-creative	4.39		4.63	4.39	4.32
Complex-simple	4.83		4.84	4.31	3.84
Important-unimportant	2.94	*	2.16	2.58	2.16

* Significant at the 5% level

extreme average overall Masculinity Index score that physics teachers assigned to physics. Physics teachers also judged physics to be less impersonal and more person-oriented, i.e. more aligned with feminine qualities, than did chemistry and biology teachers. In addition, physics teachers judged physicists to be more humanitarian than did other teachers.

(b) Physics and chemistry were viewed as being more masculine subjects by male teachers than by female teachers. When asked directly to indicate the masculinity of physics on the Characteristics of Science scale, the male teachers gave a more masculine average rating than did the female teachers. Male teachers also gave chemistry a more masculine average rating when using the Masculinity Index scale. In addition, male teachers believed more firmly than female teachers that physicists are male.

3(a) Biology teachers saw biology as being a more caring subject, i.e. possessing more feminine qualities, than did other science teachers. Specifically, biology teachers judged biology to be a warmer subject and more concerned with people. However, biology teachers also judged biology to be more logical and mathematical (masculine traits), and biologists to be more logical than did physical science teachers. Thus biology teachers' sex typing of biology was somewhat ambivalent.

(b) The effect of teacher sex upon perceptions of the gender of biology was less pervasive and less obvious than was the case for physics.

8.1.3 Sex stereotyping

8.1.3.1 Preference for Subject Characteristics

Slight differences were detected between the responses of physical science teachers and biology teachers. The physical science teachers gave higher mean ratings for girls' preferences, i.e. a more stereotyped feminine/arts reply, on 7 of the 9 characteristics judged. Although this trend is statistically non-significant, another difference between

Table 8.5 Effect of teacher sex and teaching subject upon pupils' perceived preferences for subject characteristics

A. Girls' preferences

Characteristic	Teacher sex	Mean rating from techers of Phys	Chem	Biol	Analysis of variance Significance: Teacher sex	Teaching subject	Interaction
Practical-theoretical	Male	3.77	3.50	3.35	-	-	-
	Female	4.17	2.86	3.48			
Numerical-verbal	Male	5.09	5.35	5.35	-	-	-
	Female	5.33	5.36	5.13			
Science-arts	Male	4.57	4.68	4.58	**	-	-
	Female	4.67	4.00	4.04			
Logical-intuitive	Male	4.09	4.63	4.27	-	-	-
	Female	4.50	4.71	4.30			
Masculine-feminine	Male	5.26	5.13	5.19	*	-	-
	Female	4.83	4.43	4.78			
Factual-opinionative	Male	3.51	4.00	4.19	-	-	-
	Female	3.17	3.93	3.70			
Routine-creative	Male	4.31	4.40	4.62	-	-	-
	Female	4.50	4.79	4.78			
Complex-simple	Male	4.74	4.93	4.62	-	-	-
	Female	4.50	4.50	4.43			
Important-unimportant	Male	3.20	3.03	2.92	*	-	-
	Female	2.67	2.21	2.78			
B. Boy's preferences							
Practical-theoretical	Male	2.80	3.13	2.62	-	-	-
	Female	3.00	2.57	2.63			
Numerical-verbal	Male	3.51	3.83	3.73	-	-	-
	Female	3.50	3.29	3.46			
Science-arts	Male	2.91	3.18	2.88	-	-	-
	Female	3.50	2.79	2.71			
Logical-intuitive	Male	3.31	3.30	2.81	-	*	-
	Female	3.33	3.36	2.83			
Masculine-feminine	Male	2.69	2.35	2.31	-	-	-
	Female	2.67	2.36	2.54			
Factual-opinionative	Male	3.17	3.28	2.85	-	-	*
	Female	3.83	2.57	3.08			
Routine-creative	Male	4.40	4.25	4.31	-	-	-
	Female	4.17	3.79	4.04			
Complex-simple	Male	4.43	4.50	4.15	-	-	-
	Female	4.50	3.86	4.13			
Important-unimportant	Male	2.91	2.88	2.65	-	-	-
	Female	3.00	2.29	2.79			
N	Male	35	40	26			
	Female	6	14	24			

* Significant at 5% level, ** Significant at 1% level, - Not significant

physical science and biology teachers' overall responses was found to be highly significant. Biology teachers gave lower mean ratings for boys' preferences, i.e. more stereotyped masculine/science replies, on all 9 rating scales. The probability of this result occurring by chance, as calculated from the binomial distribution, was only 0.013.

The tendency of physical science teachers to rate girls' preferences for subject characteristics more extremely (i.e. towards the stereotypically feminine/arts pole) and of biology teachers to rate boys' preferences more extremely (i.e. towards the stereotypically masculine/science pole) was also detected in the data collected from sample TSCH (see Table 8.4). It is noteworthy that the physical science teachers in this sample judged the appeal of science subjects and masculine subjects to be significantly less for girls than did the biology teachers.

The effect of teacher sex upon the replies received from sample P was discernible in the case of girls' preferences, but not for boys' preferences (Table 8.5). The female teachers indicated that girls find science subjects, masculine subjects and important subjects much more attractive than did the male teachers. This combination of subject characteristics that was rated differently by male and female teachers suggests that science teachers view the science subjects as being masculine subjects and important subjects. Not only did the female teachers indicate that science subjects and masculine subjects are more acceptable to girls by giving significantly lower mean ratings on the science-arts and masculine-feminine scales than did the male teachers, but they also tended to give lower mean ratings on all the scales when judging boys' preferences. In other words, they gave more stereotyped masculine/science replies, although the trend was not statistically significant.

8.1.3.2 Importance of Subjects

Analysis of variance revealed that all science teachers, regardless

Table 8.6 Effect of teacher sex and teaching subject upon evaluations of the importance of different subject areas

Subject area	Teacher sex	Mean rating from teachers of Phys	Chem	Biol	Analysis of variance Significance Teacher sex	Teaching subject	Inter-action
A. For male pupils							
Creative arts	Male	2.50	2.08	2.43	-	-	-
	Female		2.67	2.55			
Languages	Male	2.75	2.58	2.57	*	-	-
	Female		3.50	2.82			
Humanities	Male	2.69	2.33	2.43	*	-	-
	Female		3.00	2.73			
Science	Male	3.44	3.33	3.71	-	-	-
	Female		3.83	3.82			
Technical subjects	Male	3.44	3.50	3.57	-	-	-
	Female		3.67	3.82			
Home Economics	Male	2.25	2.08	2.00	-	-	-
	Female		2.67	2.36			
Commercial/ Business Stds.	Male	2.75	2.58	2.57	-	-	-
	Female		3.00	3.00			
B. For female pupils							
Creative arts	Male	2.81	2.42	3.00	-	-	-
	Female		3.17	2.82			
Languages	Male	2.75	2.67	2.57	**	-	-
	Female		3.50	3.18			
Humanities	Male	3.06	2.58	2.86	-	-	-
	Female		3.00	3.00			
Science	Male	2.81	3.00	3.29	*	-	-
	Female		3.33	3.64			
Technical subjects	Male	2.31	2.58	2.71	-	-	-
	Female		2.67	2.91			
Home Economics	Male	3.00	2.67	2.86	*	-	-
	Female		3.50	3.27			
Commercial/ Business Stds.	Male	2.75	3.00	2.86	*	-	-
	Female		3.50	3.36			
N	Male	16	12	7			
	Female	1	6	11			

* Significant at 5% level, ** Significant at 1% level, - Not significant

of their subject speciality, assessed a range of optional subject areas similarly, when asked to indicate the importance of qualifications in those subject areas to pupils' future lives (see Table 8.6). However, male and female teachers differed in their assessment of some of the subject areas. The female teachers tended to place greater value on qualifications in the sex-neutral and feminine subject areas than did the male teachers. Thus, the female teachers judged languages and humanities to be more important for boys. They also judged languages, home economics and commercial/business studies to be more important for girls. Most importantly, the male teachers regarded the acquisition of qualifications in the science subjects to be significantly less important for girls than did the female teachers.

8.1.3.3 Summary

1. Science teachers' ideas about pupils' preferences for subject characteristics and about the importance of subjects for boys and girls, together with their scores on the Females' Social Roles scale, clearly indicate that the sex stereotypes held by teachers vary according to their sex, and may also vary according to their principal teaching subject.

2. Physical science teachers tended to view girls' preferences for subject characteristics in a more sex stereotyped way (i.e. that girls like feminine/arts type characteristics) than did biology teachers. Furthermore in pilot studies, physical science teachers achieved lower scores on the Females' Social Roles questionnaire, i.e. they gave more traditional replies than did biology teachers. However, none of the differences on the aforementioned scales were statistically significant. In addition, the Importance of Subjects scale produced no teaching subject differences at all. Thus, those trends which were detected and which suggest that biology teachers hold less sex stereotyped views may in fact be due to teacher sex effects, since the proportion of biology

Table 8.7 Additional independent variables investigated

Variable	Masculinity Index	Females' Social Roles	Marking Exercise
TEACHER			
Age	2	4B	2C
Teaching experience (years)	2	1	3C
Status	2	2	2C
Taught compulsory science	2	1	2C
Taught in single sex school(s)	3A	2	2C
Single sex education	1	3B	2C
Social class background	3A	3B	4C
Mother's employment	1	3B	2C
CURRENT SCHOOL			
Type of school	2	2	1
Sex of school	2	2	4C
School leaving age	2	1	1
School location	2	3B	1
Background of pupils	1	3B	1

Codes

1 Not investigated
2 Investigated - no relationship
3 Investigated - possible relationship
4 Investigated - clear relationship

A Reported in Appendix 8.1
B Reported in Appendix 7.5
C Reported in Appendix 8.2

teachers who are female is very much higher than is the case in the physical science subjects.

3. Teacher sex effects upon sex stereotypes were clearly demonstrated. All three scales under discussion showed that female teachers held less sex stereotyped views than did male teachers. This was especially true when teachers were asked to consider female pupils. Of particular interest is the fact that male teachers judged science subjects to be less attractive and of less value to girls than did female teachers.

8.2 EFFECT OF OTHER INDEPENDENT VARIABLES

The effects of additional independent variables were only studied upon a single measure from each topic area, excepting attribution patterns. The measures chosen were the Masculinity Indices, Females' Social Roles scale, and the Marking Exercise. The first two scales were chosen because they were designed to measure attitudes and the formation of attitudes is known to be influenced by a person' circumstances (Triandis, 1971). The third measure was chosen because of its centrality to the present research.

A full list of the additional independent variables investigated and whether they influenced the selected measures is recorded in Table 8.7. Where definite or possible relationships are indicated, the reader is referred to the appropriate appendix for a detailed description of the findings. Table 8.7 shows that only two variables consistently affected teachers' responses on the three measures chosen for detailed investigation. They were a teacher's social class background and his/her encounters with single sex schools.

The responses of teachers from working-class backgrounds tended to be more extreme than those from teachers with middle-class backgrounds. Thus, teachers from working-class backgrounds tended to (a) sex type the physical and biological sciences more (i.e. view physics and chemistry as being more masculine, and biology as being more feminine), and (b) hold more traditionally sex stereotyped views as measured by the Females' Social Roles scale, than did teachers from a middle-class background. Furthermore, teachers from working-class backgrounds favoured the work of boys over that of girls to a significantly greater extent than did teachers from middle-class backgrounds.

The effects of single sex education as opposed to coeducational education upon teachers' responses to the different measures were more diffuse and less uniform. In general, the indications were that contact with single sex education was associated with less sex stereotyped views and less sex biased marking practices. Specifically, teachers who had received all or part of their education at single sex schools tended to obtain more liberal scores on the Females' Social Roles questionnaire than did teachers who had only attended coeducational schools. Teachers who had taught in single sex schools tended to view the masculinity of chemistry and biology differently to teachers who had only taught in coeducational schools, but the results do not form a coherent pattern. Finally, teachers who were currently teaching in single sex schools favoured the work of boys over that of girls to a significantly less degree than did teachers who were teaching in coeducational schools.

8.2.1 Conclusions

1. (Relates to Question Five, What personal and educational variables are associated with extreme views about the masculinity of science?) Two teacher variables were identified that are possibly related with a teacher's views about the masculinity of science. They are a teacher's experience of teaching in a single sex school and his/her social class background (Appendix 8.1).

2. (Relates to Question Twelve, Do science teachers from a working-class background hold more traditional attitudes towards sex roles than teachers from a middle-class background?) Replies to the short-form

Females' Social Roles scale indicated that teachers from working-class backgrounds tended to view sex roles more traditionally than did teachers from middle-class backgrounds (Appendix 7.5), but the difference between the mean scores of the two groups was not statistically significant.

3. (Relates to Question Thirteen, Do science teachers, whose mother were full-time housewives during their childhood, hold more traditional attitudes towards sex roles than teachers whose mothers were engaged in paid employment?) Replies to the short-form Females' Social Roles scale indicated that respondents whose mothers had only been housewives tended to view sex roles more liberally than respondents whose mothers had been employed (Appendix 7.5), but the difference in scores between the two groups did not reach statistical significance.

4. There were indications that certain aspects of teachers' educational experiences as pupils and their current educational experiences as teachers tended to relate with their views about sex roles (Appendix 7.5), but the relationships were not statistically significant.

5. (Relates to Question Seventeen, What personal and educational variables distinguish teachers who award very similar marks to boys and to girls from those who award very dissimilar marks?) Three variables were identified that appear to be related to the marks that teachers award to boys and to girls. They are a teacher's social class background, length of teaching experience, and the sex composition of the teacher's current school (Appendix 8.2).

8.3 RELATIONSHIPS BETWEEN THE DEPENDENT VARIABLES

8.3.1 Introduction

Attempts to identify and investigate relationships between the dependent variables were complicated and hindered by the following fa factors.

(a) The disappointing response rate to the STOSS questionnaire resulting

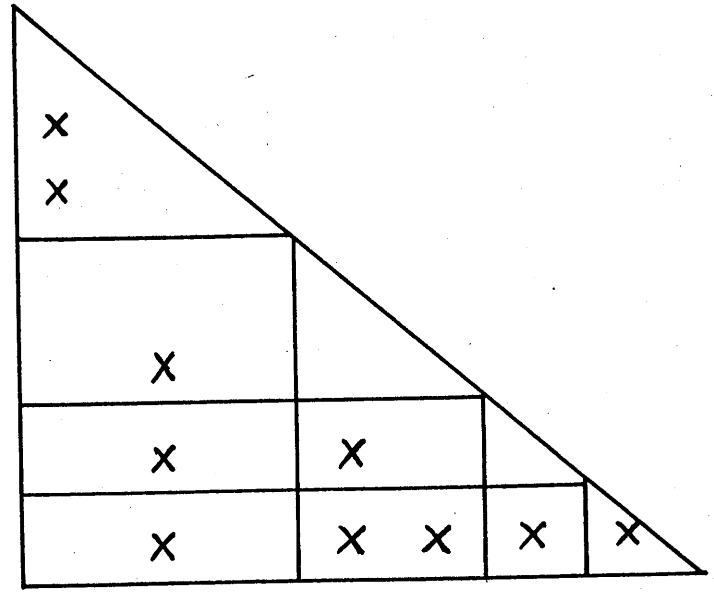
Figure 8.1 Relationships between dependent variables investigated

A. Sex typing of science
(i) Characteristics of Science
(ii) Masculinity Indices
(iii) Scientist Stereotypes

B. Sex stereotyping
(i) Females' Social Roles
(ii) Preference for Subject Characteristics

C. Attribution scales

D. Teacher Expectation/ Judgement
Marking Exercise



in a smaller than planned sample size.

(b) Generally complex experimental designs that did not facilitate inter-scale comparisons.

(c) Scales that produced data about a number of factors, rather than simply providing a single total score.

(d) The frequent need to keep the replies of physical science teachers and biology teachers separate, especially with regard to the attribution scales.

In spite of the severe limitations imposed by the above constraints, various interrelationships between the different dependent variables were studied. These investigations were guided by the interactions proposed in Chapter 1 (Figure 1.1).

8.3.2 Findings

The interrelationships between dependent variables that were investigated in detail are shown in Figure 8.1. The blocks along the diagonal of the figure refer to possible intra-topic comparisons, and the other blocks refer to possible inter-topic comparisons. It can be seen that intra-topic comparisons were only made for the sex typing of science and for teacher expectation/judgement, whereas all of the possible inter-topic combinations were investigated.

8.3.2.1 Sex typing of science - Comparisons between measures

Science teachers' perceptions of the masculinity of physics as measured by the Masculinity Index correlated positively with their replies to a direct question about the masculinity of physics that appeared in the Characteristics of Science (physics) scale ($r = 0.26$, $p < 0.001$). Comparable measurements of the gender image of biology did not correlated significantly.

Science teachers' perceptions of some of the characteristics possessed by scientists were positively correlated with their perceptions

Table 8.8 Correlations between science teachers' perceptions of scientists and the respective science subjects (N=164)

Scientist	Subject	Physicist/Physics r	p	Biologist/Biology r	p
Male	Masculine	0.34	0.001	0.07	ns
Good at maths	Mathematical	0.27	0.001	0.29	0.001
Logical	Logical	0.33	0.001	0.43	0.001
Objective	Objective	0.34	0.001	0.43	0.001
Not humanitarian	Concerned with people	-0.23	0.01	-0.32	0.001

Table 8.9 Correlations between teacher expectation (O level suitability) and teacher judgement for three standards of work (N=326)

	Teacher judgement			
Standard of work	Mark merited	Mark given	Scientific accuracy	Understanding of principles
Good	0.61	0.58	0.61	0.57
Average	0.44	0.41	0.37	0.44
Poor	0.40	0.34	0.32	0.40

of whether the characteristics were also possessed by the respective science subjects. All of the Pearson correlation coefficients recorded in Table 8.8 are significant at the 5% level (two-tailed test), with the exception of the r value for the correlation between a biologist being male and biology being masculine. Reference back to sections 7.1.1 and 7.1.2 shows that science teachers regarded biology as a not very masculine subject. Nevertheless, the teachers thought that biologists were as likely to be male as female.

In summary, the interrelationships investigated provide evidence that the scales concerned were probing closely related and concordant beliefs.

8.3.2.2 Teacher expectation/judgement - Comparisons between measures

Science teachers' expectations for the fictitious pupils considered in the marking exercise correlated positively with their judgements of the quality of the work supposedly produced by the pupils. Table 8.9 presents a number of r values that illustrate the relationship. O level suitability was chosen as the most appropriate indicator of teacher expectation for use with all three standards of work. The marks associated with each sample pair, both the marks merited and the marks given, were chosen as the most obvious indicator of teachers' assessment of the worth of the work. Scientific accuracy and understanding of principles were also included because it had been previously established that these two variables are closely related to the mark given (Appendix 7.12). Correlations between all of these different judgements of the quality of the work and O level suitability are positive and significant at the 0.001 level. The other obvious measure of teacher expectation, CSE suitability, did not produce such high correlations.

8.3.2.3 Sex typing and sex stereotyping comparisons

No link between a teacher's disposition to sex type science and to hold sex stereotyped beliefs was shown in the orginal model (see Figure 1.1). However, logic suggests that the two dispositions should correlate

Table 8.10 Correlations between teachers' perceptions of pupils' subject preferences and their views of the masculinity of the physical science subjects

	Masculinity index scores			
	Physics		Physical Science	
	r	p	r	p
Girls' preference for science subjects	-0.18	0.05	-0.25	0.01
Girls' preference for masculine subjects	-0.10	ns	-0.10	ns
Boys' preference for science subjects	0.23	0.01	0.15	ns
Boys' preference for masculine subjects	0.29	0.001	0.24	0.01

positively.

Table 8.10 indicates that correlations between masculinity indices and measures of pupils' preferences for subject characteristics were investigated. Two masculinity indices were chosen. The physics masculinity index was included since it tended to elicit the most extreme responses, and the physical science masculinity index was included as a broader measure of a teacher's tendency to sex type science. Of the various semantic differential scales used to measure teachers' views about pupils' preferences for subject characteristics, only four were considered to be suitable for inclusion in this investigation. They were the attraction of science/arts subjects and masculine/feminine subjects to girls and to boys.

The correlations described above produced the correlation coefficients shown in Table 8.10. These coefficients, although small in magnitude, indicate that teachers who tended to view the physical science subjects as masculine subjects, also tended to believe that boys like subjects that are scientific and masculine, but that girls like arts subjects.

8.3.2.4 Sex typing and attribution comparisons

No consistent interactions were identified between masculinity index scores and attribution ratings, although there were indications of relationships between scores on the physical science and biology indices and teachers' views that pupils reject or choose science because of its close association with one particular sex.

8.3.2.5 Sex typing and teacher expectation/judgement comparisons

Two masculinity indices were selected as measures of a teacher's tendency to sex type science. Obviously the chemistry masculinity index was chosen on account of the marking exercise being based on chemistry write-ups. The second masculinity index involved a composite score derived from subtracting the physics masculinity index score from the biology masculinity index score. The resulting score effectively indicated the extent to which a teacher sex typed the two extreme science subjects.

The effect of the two masculinity indices upon teacher expectation/judgement are reported in full in Appendix 8.2. Summarizing, a teacher's chemistry masculinity index score appeared not to be related to the ratings s/he awarded to boys and to girls. In contrast, a marginally significant relationship was identified between the composite score and teachers' ratings. Teachers who tended to exaggerate the masculinity of physics and the femininity of biology, were more inclined to favour the work of boys than were teachers who did not differentiate the gender images of physics and biology so sharply.

8.3.2.6 Sex stereotyping and attribution comparisons

No relationships were identified between a person's total score on the Females' Social Roles questionnaire and their attribution ratings.

8.3.2.7 Sex stereotyping and teacher expectation/judgement comparisons

The interaction between a teacher's predisposition to sex stereotype and to display sex bias in his/her assessment of pupil work was investigated via selected sex stereotyping measures and the marking exercise. Three sex stereotyping measures were chosen for their relevance and straightforward interpretation. The effect of two of these measures upon teachers' marking patterns are reported in full in Appendix 8.2. Briefly, neither a teacher's total score on the Females' Social Roles questionnaire, not the extent to which s/he differentiated between boys' and girls' liking of science subjects were related to marking patterns. The third sex stereotyping measure, the extent to which a teacher differentiated between boys' and girls' liking of masculine subjects, had to be discarded because it failed to discriminate between teachers.

8.3.2.8 Attribution and teacher expectation/judgement comparisons

No consistent relationships were identified between attribution ratings and marking patterns.

8.3.3 Summary

1. Investigations into relationships between the dependent variables were inadvertently hampered by complex experimental designs and measuring scales. Correlational relationships which were identified were characterized by small r values.

2. Comparisons between scales inquiring into the same topic area were only made in the case of the sex typing of science. The resulting correlations indicated that the scales were probing closely related and concordant beliefs.

3. Many of the inter-topic comparisons failed to evince the existence of interrelationships between the dependent variables being investigated. It was particularly notable that no relationships were found between teachers' replies to the attribution scales and their replies to the other topic areas.

4. The masculinity indices emerged as useful measures of a teacher's tendency to sex type the science subjects. Masculinity indices tended to interrelate with both sex stereotyping measures (boys' and girls' perceived preferences for science/arts and masculine/feminine subjects) and sex influenced marking patterns. Teachers who tended to exaggerate the gender connotations of the science subjects, tended also to exaggerate the attraction of scientific and masculine subjects to boys, and to overvalue the work of boys. The relationships accord with theoretical predictions.

5. No relationships were found between sex stereotyping measures and marking patterns. Comparing this result with those reported in (4) above suggests that sex typing measures are better indicators of the influence of pupil sex upon teachers' marking practices than are sex stereotyping measures.

8.3.4 Conclusions

1(a) Question Nineteen asked "Are those science teachers who regard the physical science subjects as masculine subjects, more likely to use different reasons to explain the successes and failures of boys than of girls?" This study failed to produce evidence in support of this relationship.

(b) Question Twenty Two asked "Are those science teachers with sex stereotyped perceptions and traditional sex role attitudes, more likely to use different reasons to explain the successes and failures of boys than of girls?" No evidence was produced to support this relationship.

2. Questions Twenty, Twenty One, Twenty Three and Twenty Four enquired into the existence and nature of relationships between the sex typing of science and teachers' expectations, teachers' attribution patterns and their expectations, teachers sex stereotypes and their expectations, teachers' expectations and their judgements. The analyses reported in this chapter did not directly address these relationships, since teacher expectation and teacher judgement were combined and treated as one variable. Although the original questions cannot be answered from the analyses performed, with only slight modification the questions could be examined using the available data. For instance:

(a) By combining Questions Twenty and Twenty Four attention is focused upon the relationship between the sex typing of science and teachers' expectations/judgements. It was found that teachers who tended to exaggerate the gender connotations of the science subjects also tended to overvalue the work of boys (Appendix 8.2). This finding is in agreement with the relationship specified in the original questions.

(b) Combining Questions Twenty One and Twenty Four designates the relationship between attribution patterns and teacher expectation/judgement. This study failed to produce evidence of the operation of this relationship.

(c) Combining Questions Twenty Three and Twenty Four specifies the relationship between sex stereotyping and teacher expectation/judgement. Evidence for the operation of such a relationship was not produced in this study.

3. The scale and scope of the relationships reported in this last section are disappointing. It would seem that either the scales used to measure the variables, or else the statistical analyses performed to detect relationships between the variables were inadequate. Much work still remains to be done to identify and describe the relationships that probably do exist between the topic areas investigated in this research.

CHAPTER 9

DISCUSSION

Page

9.0 Contents

9.1 SEX TYPING OF SCIENCE

9.1.1 Gender associations of science subjects

When science teachers were asked direct questions about the gender connotations of the science subjects, their replies tended to be circumspect. They maintained that both physics and biology are neutral subjects (section 7.1.4.2). However, the use of more sophisticated measuring scales readily elicited responses which indicated that science teachers do sex type the science subjects (sections 7.1.1.1 and 7.1.2). The apparent contradiction between the two sets of findings suggests that teachers unconsciously associate each science subject with one particular sex, but that when they deliberate upon the gender connotations of a subject, they rationalize that it must be or should be neutral. It is, after all, only natural that when respondents are aware of the intention and implications of a question, they will want to present themselves as reasonable, unprejudiced people.

Replies to the School Subject Characteristics scale showed that a range of school subjects are sex typed by secondary teachers. More subjects received mean ratings that placed them on the masculine side of neutral than on the feminine side. Moreover, the feminine subjects received ratings that were very close to a neutral rating, whilst the masculine subjects tended to receive clear masculine ratings. These findings suggest that there are masculine subjects and neuter subjects, but that there are few clearly feminine subjects.

Science teachers, along with teachers of other subjects, judged physics and chemistry to be masculine subjects. Both the Masculinity Index and the School Subject Characteristics scale indicated that physics is regarded as the more masculine of the two subjects. This order corresponds with indications from examination statistics. Boys constitute a higher proportion of entries for external examinations in physics than in chemistry (see Appendix 1.1). This correspondence suggests that the masculine image of a subject could be determined by the proportion of males found in the subject. Replies to the Opinions scale provide further support for this relationship (see section 7.1.4.1 for results, and below for detailed discussion). However, such a relationship is probably too simplistic. The influence of other factors, e.g. subject content, upon a subject's gender connotations are unlikely to be negligible.

Biology was judged to be a very slightly feminine subject. This perception is at odds with most of its other characteristics which identify is as a science subject, and by implication a masculine subject. Along with the other science subjects, biology was judged to be logical, factual, routine and complex. However, instead of being viewed as a numerical subject like physics and chemistry, biology was viewed as a very slightly verbal subject. This finding intimates a connection between mathematical computations and a masculine image.

The features of a science subject that contribute to its gender image were investigated more fully using the Characteristics of Science data. With regard to physics, a 'masculine' cluster was identified which contained five characteristics: masculine, technical, mechanical, mathematical, and concern with objects. These same five characteristics also appeared in the gender cluster for biology, which suggests that similar associations underlie the gender image of all the science subjects. This induction has wide ranging and important consequences. If the features of a subject that contribute to its gender image are known, then the possibility exists of altering those features with a view to modifying the gender image. For example, conscious attempts to turn chemistry into a more neuter subject might include reducing the number of calculations required, de-emphasizing the mechanics of chemistry processes and the apparatus/plant used, and instead stressing the impact of chemicals upon the environment and society. Such changes may well also alter the very nature of chemistry, but that problem need not be discussed here.

9.1.2 The views of different groups of teachers

Secondary teachers tended to hold the most sex typed views of the science subjects. Primary teachers tended to believe that science subjects are more neutral on a number of dimensions, including the masculine-feminine dimension. They rated physics less masculine and biology less feminine than did secondary teachers. This finding is particularly revealing, especially when it is coupled with the finding that primary teachers were also less stereotyped than secondary teachers in their beliefs regarding pupils' preferences for subject characteristics. Secondary teachers often assert that primary and middle schools are largely responsible for instilling ideas that physics is a boys' subject and biology is a girls' subject. Yet this research has shown that secondary teachers are most committed to the belief that (a) the physical science subjects are masculine subjects and boys are attracted to masculine subjects, (b) biology is a feminine subject and girls are attracted to feminine subjects. These findings suggest that, in reality, secondary teachers are more likely to channel pupils into sex appropriate science subjects than are primary teachers.

Non-science secondary teachers did not sex type the science subjects to a greater or lesser extent than science teachers themselves. However, non-science teachers did describe all three science subjects as being very much less creative. This finding is interesting as it suggests that either non-science teachers are unaware of the opportunities offered in modern science courses for creative thought and inventive experiments, or that science teachers place a different interpretation on creative thought and activities. If pupils, particularly girls, think of creativity in different terms to those used by science teachers, or if they fail to recognize creative elements in science courses, then they may judge science to be an uncreative subject and this perception may make science less attractive to them. Such a consequence would be particularly unfortunate if it arose from science teachers' failure to present science as a creative and inventive subject.

Among science teachers, chemistry and biology teachers tended to sex type physics to a greater degree than did physics teachers. Physics teachers viewed their subject as being significantly less masculine and more imbued with feminine values, e.g. concern with social issues, concern with people. This finding implies that physics teachers perceive physics to be less alien to girls than do other science teachers, but further work is required to test the accuracy of this inference. Biology teachers also tended to view their subject more favourably than did teachers of the other science subjects. Biology teachers tended to elevate the scientific nature of biology by thinking of it as being more logical and mathematical.

Male science teachers viewed physics and chemistry as being more masculine subjects than did female teachers. Female teachers' tendency to view physical science in less masculine terms presumably lessens any conflict that they may otherwise experience on account of their association with masculine subjects and occupations. Female teachers not only judged physics and chemistry to be less masculine, but they also tended to view biology as less feminine than male teachers. The results obtained from non-science teachers provide further support for the conclusion that female teachers tend to de-sex the science subjects.

The usefulness of biology tended to be upgraded by female science teachers. They judged biology to be significantly more relevant for both careers and family life than did male teachers. This emphasis on the usefulness of biology may arise from women's desire to improve the standing and currency value of the most feminine, and least prestigious of the three main science subjects.

9.1.3 Consequences of gender connotations

The fact that science teachers regard physics and chemistry to be masculine subjects could be indicative of a belief that science is primarily a boys' subject and therefore not entirely appropriate for girls. Several relationships between ratings on the Masculinity Indices and other scales provide some measure of support for this view. For instance, those science teachers who were most emphatic about the masculinity of the physical science subjects, were also most likely to believe that boys like subjects that are scientific and masculine, but that girls like arts subjects. In addition, teachers who tended to exaggerate the masculinity of physics and the femininity of biology, were more inclined to be biased in favour of the written work of boys, than were teachers who did not differentiate the gender images of physics and biology so sharply. These findings are clearly consistent with the theory that science teachers who view physical science as a masculine subject area are also likely to associate boy pupils with physical science, to believe that boys are better suited to the study of physical science, and to expect the work of boys to be superior to that of girls. Further research is required to confirm, clarify and amplify these interactions. Investigations along similar lines in other sex typed subject areas may also prove useful

9.1.4 Causes of science's masculine image

Science teachers expressed the opinion both overtly (their own personal stated opinion), and covertly (their perception of other people's opinions) that stereotyping and social pressures are important factors contributing to the masculine image of science. However their own opinions and their perceptions of public opinion differed superficially over the question of the most influential factors. Science teachers believe that the public principally associates the masculine image of physical science subjects with the preponderance of male scientists, whereas they themselves expressed the view that the media, e.g. advertisements, films, comics, are principally responsible for giving the physical science subjects a masculine image. But presumably the media are perceived to influence the gender image of science because of the way that scientists are portrayed, e.g. they are usually male. Thus, via their own opinions and the perceived opinions of the public, science teachers conveyed their belief that physical science has a masculine image primarily because its practitioners are male.

In their personal explanations for science's masculine image, science teachers tended to focus upon non-science and out-of-school factors. If their views are accurate, then it should be possible to change science's masculine image without changing the content of science, or the way that it is taught. Simply changing the way that science is presented to the public, i.e. its media image, could transform people's perceptions of science. However other findings in this investigation, that relate to the masculine image of science, suggest that the introduction of measures to change people's perceptions of scientists without associated alterations to the content of science would probably be ineffective. The notion that science is a masculine subject does not occur in isolation. It is closely associated with the perception that science is technical, mechanical, mathematical and concerned with objects. Since this is so, it may well be almost impossible to alter the image of science without also altering the customary practice and content of science. Certainly it would be difficult to change attitudes towards science, and resulting stereotypes and social pressures (the second influential group of factors to which science teachers attributed science's masculine image), solely by altering the portrayal of scientists. Concomitant changes within science itself would also be necessary.

9.1.5 Beliefs about scientists

When science teachers were asked how they picture a scientist, their replies clearly showed that a physicist is a very different person from a biologist. A physicist is quite probably male, whereas a biologist is just as likely to be female as male. A physicist is also more likely to display traits that are stereotypically associated with scientists. But since scientists are usually men, these same traits are also stereotypically associated with men. The fact that science teachers think of physicists as men, who display stereotyped masculine characteristics, must make it difficult for them to invisage their female pupils becoming physical scientists. This credibility problem is further compounded because male teachers are even more convinced than female teachers that physicists are male, and most physical science teachers are men.

The Scientist Stereotype scale also demonstrated the point that people tend to view themselves and their associates in a more favourable light than do other groups of people. Physics teachers indicated that physicists are more humanitarian than did the other groups of science teachers. Biology teachers rated biologists as being more logical and sociable than did the physical science teachers. Biology teachers were also more likely to categorize biologists as male. Since male dominated professions tend to enjoy greater status and monetary rewards than more sex equal professions (Touhey, 1974), this tendency may be due to biology teachers' unconscious desire to raise the status of biology. The effect upon biology teachers' aspirations for their female and male pupils is impossible to assess.

9.2 SEX STEREOTYPING

9.2.1 Written Work of Girls and Boys

The results presented in section 7.2.1.1 indicate that most science teachers recognise differences between the written work of girls and boys. Furthermore, nearly three quarters of the science teachers questioned claimed that they can generally distinguish between the written work of girls and boys. These teachers must also presumably believe that they possess the ability to surmise the sex of the author of a piece of written work. If this is the case and if teachers exercise this supposed ability, then the opportunity for pupil sex to act as a non-cognitive biasing factor is greater than in cases where the teachers are less sensitive to pupil sex. Even if teachers cannot accurately surmise the sex of the author of a piece of written work, the fact that they believe that they can is still likely to influence their perceptions of the work. So, whether teachers actually can 'sex' written work (and an experiment could easily be set up to test the accuracy of their judgements) is somewhat immaterial. Their belief that they can tell an author's sex is sufficient to trigger a host of related sex stereotyped perceptions, beliefs and expectations.

The finding that a significantly higher proportion of male science teachers (80%) than female science teachers (57%) indicated that they believe they can identify the work of girls and boys is worthy of note, in view of the fact that the majority of science teachers are male. The smaller number of female respondents supporting the notion of sex differences in written work might be due to their greater awareness of and personal concern about the principle of sexual equality.

Nevertheless, the views revealed by this study indicate that either substantial differences do exist between the work of boys and girls, or that popular cultural stereotypes concerning the characteristics of boys and girls and their work are accepted by most teachers. This suggestion that most teachers associate particular characteristics with boys' and girls' work is supported by the writings and findings of other workers (Spender, 1977). Davies and Meighan (1975) reported that "Fields where the girls were perceived to excel nearly all related to their greater devotion to work: they were more conscientious, precise, organized and better at written work" (p.174). Some studies have attempted to explore these opinions in greater depth in order to determine their basis. However, few areas have been so investigated, and thus the focus of the following discussion is largely determined by the availability of pertinent references.

The opinion most commonly mentioned was that girls are more concerned with orderliness and neatness than boys. Gardner (1974) obtained information on this trait when he administered the Personal Preference Index to Australian high school pupils. The data indicated that girls are more preoccupied with orderliness. The greater attention paid by girls to the presentation and fine details of their written work means that their work is indeed neater than that of boys (Smail & Kelly, 1984).

Several explanations have been proposed to account for girls' greater concern about the appearance of their work. Samuel (1981) asks whether 'neatness' from girls is a result of their being brought up to try to please people. Several of the teachers' replies indicated that they thought this to be the case. A male chemistry teacher wrote that girls give "a general impression of a wish to please", and a male religious education teacher wrote that girls "tend to try to please teacher". By linking this idea with other ideas discussed in section 9.4.3, it could be argued that girls produce neat work in the belief that it will please the teacher, be viewed more favourably, and thus be awarded higher marks. Some writers suggest that girls desire high marks since they signify social approval. Ebbutt (1981b) suggests that girls produce neat books in order to compensate for the lack of other more tangible end products from science lessons. This suggestion could help to account for biology's attraction to girls, since biology has traditionally entailed the production of pages of detailed diagrams. However, the HMI publication 'Girls and Science' (DES, 1980) suggests that although girls are well able to produce exceedingly neat written work and diagrams, their competence should not be equated with enthusiasm for this passive style of working and learning.

Girls are not only neater in their work, but they are also expected to be neater. Sharpe (1976) wrote "Girls are expected to be more tidy in their ways and in their work, neater in handwriting and the presentation of material. Boys however could get away with messy books and untidy behaviour" (p.152). Borrowing an idea from Rosen (1972), Sharpe suggested that the stressing of neatness may be of questionable educational value to girls. "This is a subtle form of restriction of expression, and may develop into an obsession with form rather than content" (p.152). This idea was also expressed by several of the respondents. For example, one teacher stated that girls "stress presentation rather than content". A female maths teacher wrote about girls' work, "Too much concentration on the drawing of a diagram; little thought given to the mathematical importance of the diagram". The most emphatically stated view on the form/content balance of girls' written work came from a male biology teacher. He wrote that girls are "sometimes more concerned about appearance than correctness". In contrast, comments referring to the written work of boys, stressed boys' preoccupation with content at the expense of presentation. For example, a male chemistry teacher wrote that boys "concentrate more on content than appearance". There was a remarkable degree of agreement over the form/content balance of boys' and girls' written work. This seems to indicate either that it does differ between the sexes, in which case the part played by schools in bringing about this difference should be investigated carefully, or that the respondents have readily accepted a common stereotype. If the latter is the case, there is always the possibility that teacher expectancy effects could come into operation.

Of the various features mentioned by the subjects, those relating to verbal ability have been most extensively researched and documented. There is now substantial evidence showing that females are superior on tasks of grammar, spelling and word fluency (reviewed by Garai & Scheinfeld, 1968; Maccoby & Jacklin, 1975). In view of the wide acceptance of this generalization that girls have greater verbal ability than boys, it is surprising that so few mentions were made in this category.

Associated with girls' better verbal ability is their greater facility to write at length. This characteristic, which was mentioned by the subjects, has also been established by controlled investigations (Labrant, 1933). It is worth noting that some teachers associate length with irrelevance and brevity with precision. A male physics teacher noted that girls produce work that is "wordy, often with unnecessary detail". Another male physics teacher wrote that girls' work is "uncritical, including unnecessary detail, not appreciating the essentials". In contrast, boys "seem on the whole to be more able to grasp overall ideas, despite inattention to minute detail". A male integrated science teacher explained further the perceived superiority of boys' written work. "Broad outlines sound, detail sometimes lacking or garbled. Whole greater than the sum of the parts." It does seem that girls' greater ability to write at length can lower the perceived standard of a piece of work, rather than raise it. Perhaps this is an example of a male characteristic being valued over the female one.

The findings reported in section 7.2.1.2, together with the foregoing discussion, suggest that teachers tend to devalue or discount the work characteristics of girls. The written work of girls is principally described as being 'neat', 'well presented' and 'thorough'. Girls obviously 'try hard' and are 'conscientious'. Although all these valuations are positive, they refer to behavioural traits rather than to mental ones. By focusing upon the conscientiousness of girls, teachers seem to overlook or devalue the actual content of girls' work. The reverse tends to happen when boys' work is considered. Teachers prefer to focus upon boys' innate cognitive abilities and to excuse their inferior behavioural traits.

Overall, the respondents showed remarkable agreement over the commonly perceived features which characterize the written work of boys and girls. A brief review of the findings of selected experimental investigations has revealed that certain of the features mentioned correspond to established sex differences between boys and girls. However, other supposedly distinguishing features have not been verified by research and thus they are more likely to reflect popular stereotypes concerning the characteristics of boys' and girls' work. It would appear that teachers' perceptions are considerably influenced by such stereotypes. So too, presumably, are teachers' reactions to their pupils' work and their judgement of the work of each sex.

9.2.2 Preference for Subject Characteristics

The results presented in section 7.2.2.1 show that science teachers believe that boys' and girls' preferences for subject characteristics differ significantly. Most noticeably, boys are thought to prefer subjects that can be described as numerical, science, logical and masculine; whereas girls are thought to prefer subjects that are verbal, arts, intuitive and feminine. In addition, practical and factual subjects are considered to be more attractive to boys, whilst simple and creative subjects are more attractive to girls. It would appear that teachers have few reservations about expressing such beliefs. In both the STOSS and COSS questionnaires, the two scales that enquired into boys' and girls' preferences for particular subject characteristics appeared immediately beneath each other. Thus it was quite obvious that the replies for boys and girls would be compared. The sex differentiated responses received from the teachers under such circumstances suggest that they considered their opinions to be both accurate and in accord with other teachers' views.

The beliefs expressed by the science teachers correspond with findings and assumptions recorded in the literature. For instance, Whyld (1980) wrote that boys are often assumed to be more rational, objective and logical; several reviews have shown that boys are more interested than girls in the physical sciences (Gardner, 1974; Ormerod & Duckworth, 1975); Hutchings (1967) reported that some girls dislike the factual nature of science; and it is often stated that girls dislike and are deterred by difficult subjects (Kelly, 1979). Gardner (1974) found that girls' performance on numerical tests was weaker than that of boys, and DES (1980) reported that many girls encounter difficulties with mathematics. Kelly (1979) suggested that girls' rejection of science stems from the masculine image of science. Girls want to be involved in subjects that convey feminine values. Evidence concerning the practical-theoretical dimension is more contradictory. Kelly (1976a) suggested that many people believe that girls cannot do practical subjects, but reports from girls themselves indicate that they enjoy practical work (Curran, 1980; DES, 1980). The teachers in this investigation probably underestimated the attraction of practical subjects to girls. On the other hand, they seem to have overestimated the appeal of theoretical subjects to girls, for the literature indicates that girls often encounter difficulties with abstract and theoretical subjects (Curran, 1980; DES, 1980).

Girls superior verbal ability is widely accepted (Maccoby & Jacklin, 1975). It presumably accounts for girls liking of subjects in which they can express their own opinions. Girls also like to use their imagination in their work and to be creative (Kelly, 1976a). These expressive-verbal aptitudes and preferences of girls are clearly identifiable in the teachers' replies in this study. Even more interesting are the results of the factor analysis which showed that the teachers linked the two characteristics, creative and verbal, when thinking about the subject characteristics preferred by girls. This association is quite plausible, given the evidence above. Perhaps it is surprising that opinionative and intuitive did not load on the same factor as well. The fact that boys' perceived preference for creativity received a very neutral rating, and was not associated with any of the other characteristics, could mean that the teachers regarded it to be an inappropriate variable when describing the types of subjects that are attractive to boys.

When the teachers' replies are compared with pupils' replies it can be seen that girls' preferences for subject characteristics are actually closer to the characteristics that boys prefer than teachers believe. The discrepancies between teachers' opinions about boys' preferences for different subject characteristics and boys' actual preferences are more variable. However, since the teachers tended to give more extreme ratings to both girls' preferences and to boys' preferences than the pupils did themselves, they were in fact magnifying the difference between boys' and girls' preferences. This finding suggests that perhaps the teachers' beliefs were being influenced more by commonly accepted sex stereotypes than by their objective knowledge of pupils likes and dislikes.

Teachers' perceptions of the subject characteristics preferrred by boys and girls may also be influenced by the occupational roles that they associate with men and women. If science teachers regard scientific research to be a male occupation (which they do, according to the results in section 7.1.5), then presumably they expect boys to be more interested in science subjects than girls. Similar links can be postulated between boys' perceived interest in numerical subjects and the preponderance of males in mathematically-based occupations, and between boys' preference for important subjects and men's control of high status occupations. Finally, the custom of using the terms 'masculine' and 'feminine' to describe school subjects may partly have arisen because people link certain school subjects with sex typed occupations.

Science teachers seem to link the two variables 'feminine' and 'arts' together when describing subject characteristics preferred by girls, but do not make a similar linkage between the variables 'masculine' and 'science' when describing subject characteristics preferred by boys. This suggests that science teachers believe that a subject has to be both 'feminine' and 'arts' in order to appeal to girls, whereas boys are attracted to 'science' subjects whether or not the subjects are also 'masculine'. The separation of the two characteristics 'masculine' and 'science' in the teachers' minds means that a large number of school subjects are potentially attractive to boys. In contrast, teachers presumably believe that girls are attracted to a much smaller range of school subjects, since ideally they should be both 'feminine' and 'arts' subjects.

Striking similarities exist between science teachers' perceptions of the characteristics of the science subjects and their beliefs about the subject characteristics preferred by boys. The physical science subjects are described by science teachers as being numerical, science, logical, masculine, factual, routine, complex and important (section 7.1.1.1). All of these characteristics are believed to appeal to boys more than to girls (section 7.2.2.1). Thus in science teachers' minds, there is a quite close match between the characteristics of physical science and the subject characteristics preferred by boys. In contrast, there is a serious mismatch regarding girls' perceived preferences. Thus, not only do science teachers judge that arts subjects have more appeal for girls, but they also believe that the characteristics of physical science are generally unattractive to girls. The characteristics of biology are judged to be less extreme than those of the physical science subjects, and so biology matches with girls' preferences more closely. By implication, biology must be viewed as a more acceptable science subject for girls.

The tendency of science teachers to exaggerate girls' preference for feminine related characteristics and boys' preference for masculine related characteristics means that they perceive girls to be farther removed from science than they actually are, but that they magnify the affinity between boys and science. It is noteworthy that male science teachers tend to exaggerate girls' preference for feminine related characteristics even more than female science teachers. On the science-arts and masculine-feminine scales, the male teachers gave significantly different replies to the female teachers. Since the majority of science teachers are male, it is unfortunate that, as a group, they are least sure about the attraction of science to girls. The greater willingness of female science teachers to acknowledge that girls can be attracted to science presumably stems from their own circumstances.

Science teachers who believe that girls are not attracted to science, may anticipate that girls will drop science when the opportunity arises. Furthermore, they may fail to encourage girls to continue studying science because of their belief that girls prefer arts subjects. These, and other related beliefs that have also been investigated in this study, may have wide ranging and deleterious effects upon girls' position in science. For example, teachers may hold different expectations for girls studying science than for boys, and may behave differently towards girls than boys, with the result that boys may receive more attention and teaching.

The mismatch in teachers' minds between girls and science could be lessened by shifting teachers' views about both girls' preferences and the characteristics of science. There is an obvious need to de-stereotype teachers' views about girls and their preferences. Particular attention needs to be paid to the beliefs of male science teachers. In addition, teachers should be encouraged to think of physical science as a more open subject area, i.e. a more neutral and less masculine subject area.

9.2.3 Females' Social Roles

In 1980 Delamont wrote "The attitudes (towards sex roles) of both male and female teachers either in training or in the occupation are unknown". This investigation has provided many details about practising science teachers' views of sex roles. The results of the main study showed that 40% of the teachers held very traditional, sex stereotyped attitudes about women's social roles. But the fact that nearly a quarter of the respondents expressed very non-stereotyped attitudes should not be overlooked. Unfortunately, the scores obtained by the science teachers cannot be compared with those of other groups of the population as the shortened form of the scale used has not yet been administered to other samples. However, the two long scales used in the pilot studies have been used with other samples and so comparisons can be made. Slade and Jenner (1978) recorded that female lecturers and research workers achieved a mean score of 80 on the attitude scale used in the first pilot, female university students scored 75, and housewives scored 53. In this research, female science teachers obtained a mean score of 78 and men scored 66. These scores suggest that female science teachers hold similar attitudes to other educated, professional female groups. Although the attitudes of male science teachers are more traditional, they seemingly do not compare too unfavourably with those of different female groups. However, it should be born in mind that the mean score of the group of housewives is considerably below a mid-score (62.5) and thus reflects quite traditional beliefs. Normative scores for the scale used in the second pilot have been provided by Singleton & Christiansen (1977). They recorded a mean score for female students of 77 and for male students of 66. The science teachers' mean scores of 76 for women, and 69 for men are surprisingly similar considering that the student samples were younger (median age 25) and a different nationality (American).

The finding in both of the pilot studies as well as the main study that men had significantly more traditional sex role attitudes than women is consistent with the literature (Jean & Reynolds, 1980; Singleton & Christiansen, 1977; Spence et al., 1973). These different attitudes presumably reflect different outlooks on life. Williams (1977) has suggested that it is in men's self-interest to encourage women to conform to traditional sex roles. Such a strategy protects men's jobs and status, and enables them to maintain their dominance in the work place and also in the home.

The very high scores achieved by some of the female science teachers indicates that they held very liberal attitudes towards females' social roles. Presumably their non-traditional sex role attitudes were one of the factors that allowed them to consider studying science, a masculine subject, and enabled them to proceed and succeed in a male dominated field.

The finding from the second pilot and the main study that teachers over 40 held significantly more traditional sex role attitudes than did teachers under 40 is consistent with previous research (Jean & Reynolds, 1980; Spence et al., 1973). The more liberal sex role attitudes of younger people is presumably linked with the gradual change in adult sex roles, particularly the broadening of females' roles, that has occurred over the last couple of decades (Evans, 1982; Mason et al., 1976).

The traditional sex role perceptions of a sizeable minority of the science teachers questioned are no longer adequate reflections of present-day reality. In 1981, 63% of all women between the ages of 16 - 59 were economically active, i.e. employed, self-employed or unemployed (EOC, 1984). The economic activity rate among married women with dependent children was 49%, and among married women with child(ren) under 5 years the activity rate was 25% (EOC, 1984). These figures clearly indicate that large numbers of women are to be found in the work place, and consequently that women's place can no longer be solely in the home. Furthermore, since considerable numbers of women who have children are working, the responsibility of looking after the children must be shared with other people. This is particularly true in the case of single parent families, if the parent is employed. The incidence of single parent families has increased considerably over the last two decades (CSO, 1983). It is estimated that currently one in eight families with dependent children has only one parent (NCH, 1984). Nearly 90% of these single parent families are headed by women (EOC, 1984). If account is also taken of the fact that large numbers of single women without dependent children (single, widowed, divorced, separated women) have to fend for themselves, then it becomes obvious that many women are not working merely to earn pin money, but to earn essential bread money.

Society is changing rapidly. Teachers need to be aware of current trends and to prepare their pupils to cope with likely circumstances in their future lives. If teachers cling to traditional sex role stereotypes, e.g. a women's place is in the home, they are likely to convey ideas that contradict the experiences of many of their pupils and conflict with the expectations and aspirations of other pupils.

The educational system is one of society's principal institutions for the socialization of children. The attitudes and behaviour of teachers is crucially important in this process. They help to provide youngsters with "a reflection of society's expectation of their lives, how they will be valued, and what they may become" (McCune & Matthews, 1975, p.296). Unfortunately, teachers who hold traditional sex role beliefs may unwittingly be transmitting unrealistically conservative role expectations, instead of broadening girls' horizons.

Teachers' beliefs about women's social roles will also influence their views about girls' education. The belief that women's lives should revolve around the home rather than around economic or community tasks, must influence the perceived importance of different subjects to girls' future lives. Presumably subjects, and topics within subjects, that are judged to be important for boys' future lives as wage earners may not be viewed as being so important for girls' future lives as housewives and mothers. This tendency to sex stereotype subjects has actually been identified in this study (see the next section).

Attempts to improve girls' participation in non-traditional subject areas (including physical science), to eliminate some of the worst forms of sexism encountered in schools, and to provide equal educational opportunities for girls and for boys must be seriously hampered by the traditional sex role beliefs held by some teachers. For progress towards greater educational equity for girls and boys to proceed, steps must be taken to make teachers more aware of their attitudes and the possible effects of those attitudes. In the broad field of psychology, following the publication, in 1970, of Broverman et al.'s article which drew attention to sex role stereotypes amongst mental health practitioners, considerable publicity and effort was devoted to the question of sex role stereotyping and sex bias in allied professions (APA, 1975). The topic of sex role biases was introduced into training programmes, together with ways of eliminating such biases (Harmon et al., 1978). Later research has indicated that these approaches have succeeded in liberalizing the sex role stereotypes held by psychologists (Tetenbaum et al., 1981). A similar strategy could be adopted to improve teachers' sex role attitudes.

9.2.4 Importance of Subjects

Teachers believe that science and technical subjects are of greater importance to boys than to girls, whereas home economics and commercial/business studies are more important to girls. Moreover, teachers are so committed to these beliefs that they are prepared to express them quite openly. It must be remembered that each teacher was asked to rate the importance of a range of optional subject areas for both boys and girls. Since the two scales appeared immediately beneath each other, it was quite obvious that the replies for boys and girls would be compared.

Science teachers' views about the importance of different subject areas are more sex differentiated than are those of teachers of other subjects. This conclusion has also been reached in a major research study conducted by Pratt, Bloomfield and Seale (Pratt, 1984; Seale et al., 1982). Their work neatly complements and broadly supports the findings of this investigation.

To date, the literature has carried conflicting reports of the generality of science teachers' concern over the position of girls in science. Her Majesty's Inspectorate (DES, 1980) investigated 15 selected schools and found that most of the schools were concerned about girls' low uptake of the physical science subjects. In contrast, Whyte & Smail (1982) report that at the beginning of the GIST project involving 10 Manchester schools, the majority of the teachers were not concerned about girls' rejection of physical science and technical crafts subjects. They regarded the girls' actions to be a perfectly natural feature of school life, and actually anticipated such a response when planning timetables and option systems. The quotes recorded in Chapter 5 provide further support for the view that science teachers regard physical science to be of particular importance for boys, and that they are not overconcerned about retaining girls in the subject. The results from the Importance of Subjects scale would seem to support the contention that science teachers place less emphasis on keeping girls in science than on keeping boys. If science teachers believe that science subjects are much more important for boys than for girls, then they are likely to positively encourage boys to continue their science subjects, but not to be too perturbed when girls decide to drop science.

It is particularly worrying that science teachers judge the science subjects to be less important for girls than do teachers of other subjects. Not only are science teachers underestimating the importance of science for girls compared with other teachers, but they are probably also underestimating the importance of science for girls compared with the girls' parents. Kelly et al. (1982) asked parents how important they thought it was for their child to continue various subjects when they became optional. Of the traditional school subjects, they indicated that English and maths were the most important subjects for girls, followed by physics. Physics was judged to be even more important for girls than for boys. These findings suggest that parent may hold more enlightened views regarding the importance of non-traditional subject choices to girls' educational experience and career prospects than do teachers.

The findings regarding science teachers' views about the comaprative importance of science for boys and for girls closely intermesh with findings from other scales. For example, the underestimation of the importance of qualifications in science and technical subjects to girls in their future lives is presumably linked with the belief that a girl's career is relatively unimportant, since she will soon marry and then devote her time to caring for her husband and children. Again, if science qualifications are believed to be more important for boys, then presumably science teachers hope and expect that more boys will take external examinations in the science subjects. Support for this argument was provided by the Marking Exercise, in which 'boy' authors were judged to be significantly more suited for O level physical science courses than were identical 'girl' authors. Science's masculine image also links in with these findings. The three views that (a) science has a masculine image, (b) science is more important for boys, and (c) boys are better at science, are complementary and most probably mutually reinforcing.

The findings discussed above indicate that there is a need to work with science teachers to raise their awareness of the importance of qualifications in science to girls. The finding that male science teachers regard the acquistion of qualifications in the science subjects to be significantly less important for girls than do female teachers, suggests that special attention should be devoted to improving the attitudes and beliefs of male science teachers. The fact that male science teachers generally devalue feminine subjects compared to female science teachers, adds further support to the foregoing recommendation.

9.3 ATTRIBUTION PATTERNS

9.3.1 Reasons for Success/Failure at Science

The results of the main study, together with those from the two pilot studies and the interviews, provide considerable information about the causal factors used by science teachers to explain the success and failure of pupils in the science subjects.

Early in the investigation, it was established that science teachers do not use luck in their attributions. In response to the open-ended questions asked during the interviews, not one of the 20 teachers mentioned luck as a factor contributing to academic success or failure in the science subjects. Luck was included in the first pilot because of its central position in the theoretical typology proposed by Weiner et al. (1971). However, the respondents indicated that they believed luck to be the least important of the factors under consideration. In fact, none of the teachers indicated that luck contributes to success, and only one teacher believed that luck made any contribution to failure.

The relative unimportance of luck in teachers' attributions has also been found in other studies. Both Bar-Tal and Guttmann (1981) and Lorenz (1982) found that teachers made less use of luck than any other causal factor in explaining academic performance in maths. Some researchers have not even included luck as a category in their attribution work with teachers (Burger et al., 1982; Cooper & Burger, 1980). The under-use of luck has also been detected in other groups of respondents (McHugh et al., 1982). Such research evidence casts doubt upon the centrality and importance of luck as an attribution factor. In turn, the interpretation of sex differences in the use of luck as a causal attribution becomes problematical if luck is not a conceptually significant factor. It would appear that the concept of luck needs to be better clarified and defined. It could be that luck has different connotations to different groups of people and in different circumstances. More research is required to ascertain the empirical, as opposed to the theoretical, importance of luck in various contexts.

The comparative importance of the different causal factors investigated in the main study was determined by ranking their mean ratings. It transpired that the six variables heading the list of factors contributing to success in science all referrred to personal attributes or behaviours displayed by the pupils, i.e. internal causes. Thus the teachers' ratings indicated that they believed that pupils themselves were responsible for their own success. Effort was considered to be the factor which contributes most to success in science. Ability was ranked third. This position was lower than expected, considering that ability had been the most commonly mentioned variable in the interviews.

Personal attributes and behaviours also headed the list of factors that teachers believe contribute most to failure in science. Such results indicate that teachers place most of the blame for failure upon the pupils themselves. Lack of effort was considered to be the most important factor that contributes to failure, followed by affective factors, then lack of attention. Lack of ability was ranked fifth. Again, compared to the interviews when limited intellect was mentioned more frequently than any other explanation, the ranking of ability is surprisingly low.

Some evidence relating to the hypotheses that boys' success in science would be attributed to stable internal factors, such as ability, whilst girls' success would be explained in terms of unstable factors, such as effort, was supplied by the analysis of variance results. The second-order interactions between pupil sex and teaching subject for ability and effort provided a small measure of support for the hypotheses. Slight inferential support also appeared in the cluster analysis. Two of the three clusters formed by the attributions for girls consisted of external factors. In contrast, only one of the four clusters for boys consisted entirely of external factors. Thus the dimensions underlying boys' and girls' success in science have a different emphasis. External factors, over which pupils have no control, pervade explanations of girls' success more than boys' success. Therefore, the success of boys must rest more upon internal factors.

None of the analyses conducted with the attributions for failure supported the hypotheses regarding sex differences between the factors used to explain pupil failure in science.

Turning to individual external factors, the teachers considered standard of teaching to be a more important factor contributing to both success and failure than family support. Thus the teachers acknowledged their efficacy in the learning process. They believed that the academic achievements of pupils depends more upon their efforts than those of the parents. Similar findings were also obtained by Bar-Tal and Guttmann (1981).

More detailed analysis of the ratings given to family support revealed that female teachers considered the variable to be more important to both success and failure in science than did male teachers. Since women usually devote more time, energy and interest to home and family matters, it is consistent that they place a higher value on the families' capacity to influence children's academic achievements. More surprising was the finding that family support is considered to contribute more to the success and failure of boys in science than of girls. It is often stated that external factors, e.g. the attitudes and behaviour of influential people, are more likely to influence the level of girls' performance than of boys' (Mischel, 1967). Thus theoretically one would expect level of family support to contribute most to girls' academic achievement. The fact that the teachers believed that family support exerts a greater influence upon the performance of boys could either mean that teachers think that boys are more impressionable than girls, or that teachers think that parents provide more extreme levels of support for their sons than for their daughters.

Out-of-class experience received low rankings indicating that the teachers believed that it contributes little to success or failure in science. This view is not shared by some educational researchers. Research has shown that boys are more likely to have gained experience of mechanical and spatial activities outside of the formal educational system than are girls (Smail & Kelly, in press; Smail et al., 1982). As a consequence, when boys enter secondary school they obtain higher scores than girls on tests of spatial visualisation (Smail & Kelly, 1984). Since spatial ability is believed to be a relevant aptitude for studying science (Kelly, 1976a), it is argued that boys' greater competence in this area may facilitate their performance in science and technical craft subjects (Kelly et al., 1981). In addition, it is often suggested that because girls lack prior experience of technical and science-related activities, they will lack confidence when engaged in practical work and will be unduly anxious about their performance in science subjects (DES, 1980; Kelly et al., 1981). Research has shown that self-confidence is necessary for achievement in many academic fields (Fox et al., 1979) and that anxiety hinders creative thought (Maccoby, 1976). The relationship between self-confidence and achievement in maths is well documented (Fennema & Sherman, 1977). Kelly (1982) maintains that lack of self-confidence and a fear that science is too difficult contribute to impairing girls' achievement in science and detering them from continuing their science studies.

The findings from the present study that the consensus of a large number of practising teachers was that out- of-class experience contributes little to success or failure in science is difficult to reconcile with theoretical stances. Either many educational researchers are overemphasizing the importance of out-of-class activities to science achievement or science teachers do not appreciate the direct and indirect contribution that out-of-class experience can make to achievement in science. Analysis of variance results would seem to refute the latter explanation. Principal teaching subject produced a main effect for out-of-class experience. The teachers acknowledged that out-of-class experience contributes more to success in physics and biology than in chemistry. Intuitively this finding seems plausible. If teachers differentiate between the contribution of out-of-class experience to success in different subjects, then they cannot be altogether unaware of the potential contribution this factor can make. Thus, the findings suggest that practising teachers may well be better judges of the importance of out-of-class experience as a causal factor to success and failure in science than academics who have little contact with children learning science in schools.

The teachers did not consider that success in the science subjects was due to their being easy subjects. However, subject difficulty was believed to be a fairly important factor contributing to failure in science. Indeed, the difference in rank of subject difficulty as a contributory factor to success and failure was greater than for any other factor. These findings clearly show that the teachers considered the science subjects to be difficult subjects. However, the difficulty of the individual science subjects was not thought to contribute equally to failure in those subjects. Subject difficulty was judged to be most important to failure in chemistry but least important to failure in biology. Thus biology was viewed as the easiest science subject.

Assistance from peers was ranked the least important factor contributing to success in science. This finding clearly indicates that the study of science is a competitive activity and not a cooperative activity. It has been suggested elsewhere that physical scientists are very competitive in their work (Mitroff et al., 1977). It would seem that the foundations of this competitive approach to work are laid during school science lessons.

Distraction by peers was ranked fairly low as a factor contributing to failure in science. Such a result implies that discipline is fairly tight in most science lessons and that there is minimal interaction between pupils. It is interesting to note that there were differences between the three science subjects. Distraction by peers was least important to failure in physics and most important in biology. This finding suggests that there is least interaction between pupils in physics and most in biology. Perhaps biology teachers encourage greater cooperation between pupils. If this is the case, since a cooperative approach to tasks is considered to be an essentially feminine approach, it could be related to the fact that school biology has a feminine image.

Finally, it is worth mentioning why so few comparisons have been made in this discussion with the findings of other workers. Two particularly relevant studies have been reported by Bar-tal and Guttmann (1981) and Burger et al. (1982). However, neither of the two studies was conducted in the United Kingdom and both studies employed small samples (N=8, Bar-Tal & Guttmann) of female primary teachers. Therefore, their detailed findings are unlikely to be repeated in the present study, because of differences in nationality, sex, teaching subject and teaching level between the samples. Furthermore, small samples are more likely to produce values that are farther from the true mean of the population. For these reasons, it seemed inappropriate to make more than passing reference to the results obtained by Bar-Tal and Guttmann and Burger et al.

9.3.2 Reasons for Choosing/Dropping Science

Besides providing useful descriptions of science teachers' views of the reasons why pupils choose to continue with the science subjects or to drop them when their study becomes optional, the results of the main study also provide revealing insights into the teachers' perceptions of the characteristics of the subjects that they teach, and the motivating forces behind male and female pupils' subject choices.

At the descriptive level, the teachers' views about the comparative influence of each factor upon subject choice is interesting, because of the scarcity of published reports upon this topic. Using a case study approach, Ebbutt (1981a) interviewed the seven science teachers at a girls' grammar school and asked them why they thought girls opted as they did in that school. The teachers mentioned similar factors to those investigated in the present study, but because of the small sample used in Ebbutt's study, quantitative comparisons cannot be made. Large scale investigations into science teachers' opinions regarding the motivating factors behind pupils' subject options appear to be lacking. It is hoped that the findings of the present study will fill this gap in our knowledge and understanding of science teachers' thinking. Moreover, by comparing the teachers' views with published reports of the reasons that pupils give for choosing or dropping subjects, it is possible to assess the accuracy of the teachers' views.

The teachers' perception that pupils primarily choose or reject a subject because of their liking for the subject and because of the subject's relevance to their future career plans is well supported by the literature (Dickson, 1979; Kelly, 1981; Reid et al., 1974). There is also evidence that pupils drop science subjects because they find the subjects, particularly the physical sciences, too difficult (Gannon, 1980; Kelly, 1981). However, compared with pupils' assessment of the comparative importance of the different contributory factors, the teachers seem to have overestimated the importance of subject difficulty as a deterring factor.

Many teachers were of the opinion that pupils' choice of science subjects is often influenced by their teachers. Teacher influence can operate directly through liking of the teacher, and indirectly through liking of the style of teaching. Both aspects of teacher influence were believed to be important explanatory factors. The science teachers interviewed by Ebbutt (1981a) also expressed the opinion that teaching style and the teacher influence subject choice. Furthermore, teachers of other subjects believe that pupils choose subjects because they like the teachers, although ideally teacher popularity should not be an important factor (Reid et al., 1974). In the same study, when the teachers' pupils were questioned, they rated liking of teacher very low on their list of reasons for choosing subjects. Thus the teachers had apparently overrated their influence upon their pupils' subject choice. The same phenomenom seems to have occurred in the present study.

The teachers' assessment of the comparative importance of teacher influence to boys and to girls also differs from pupils' reports. Kelly (1976a) refers to a number of studies which showed that girls were more influenced by their science teachers than were boys. However, the teachers in this study were of the opinion that boys are more likely to drop a science subject because they dislike the teacher than are girls.

The teachers' view that parents exert considerable influence over their children's subject choices is prevalent in the literature. Ebbutt (1981a) and Reid et al. (1974) report that the teachers they contacted thought that parental infuence was an important factor in pupils' subject choices. In some studies, pupils also acknowledge their parents' influence (Ormerod, 1981), but in others they place parental influence low on their list of reasons for choosing a subject (Reid et al., 1974). Davies and Kandel (1981) suggest that differences in pupils' perceptions may arise because there is a tendency for pupils to underestimate the influence of significant others. If this is the case, then the teachers' assessment of the importance of parental influence in the present study may well be reasonably accurate.

The findings of this study support the commonly held view that the gender image of the science subjects differentially influences the uptake of the sciences by boys and girls. It is often suggested that girls reject the physical sciences because of their masculine image (Curran, 1980; Kelly, 1981). The analysis of variance results obtained in this study clearly indicate that the teachers believe that some girls reject the physical science subjects and that some boys choose these subjects because they are considered to be boys' subjects. In addition, the teachers indicated a tendency for girls to choose and for boys to reject biology because it is a girls' subject.

One reason why the physical sciences are commonly viewed as boys' subjects is that boys are generally the majority sex. Some writers have suggested that the prospect of minority status deters some girls from choosing the physical sciences (Scott, 1980; Seale et al., 1982). The teachers questioned in this study thought that pupils' choice of science subjects was little influenced by the likely composition of the teaching group. This suggests that pupils may be less worried about being in a minority position than educationalists think, or that teachers fail to detect pupils' worries about this problem, or that the teachers failed to grasp the full implications of the question. Without probing, it is impossible to judge which is the best explanation.

The results of analysing the teachers' replies reveal much about the teachers' perceptions of the characteristics of the subjects that they teach. They think of the science subjects as difficult subjects, since they believe that more pupils drop the science subjects because they find them difficult than for any other reason. Consistently, they believe that few pupils choose science subjects because they find them easy. However, science teachers do not think that all the science subjects are equally difficult. The analysis of variance results revealed that biology is seen as being less difficult than physics and chemistry.

When the teachers replies were subjected to cluster analysis, the perceived difficulty of a subject was always grouped with whether the pupil liked the content of the subject. This was true for both boys and girls. The analysis of variance results provide further evidence of a link between liking for a subject and finding it easy. It has already been pointed out that the teachers considered biology to be the easiest science subject. They also believe that biology is more likely to be chosen and less likely to be dropped because of its content, than physics or chemistry. This difference may be partly due to differences in the difficulty of the physical and biological sciences, but it could also be accounted for in terms of biology syllabuses being more appealling to the majority of pupils than physics or chemistry syllabuses. The implication seems to be that the physical sciences could be made more attractive to more pupils by simplifying and modifying syllabuses.

Teachers believe that pupils are most likely to drop physics and least likely to drop biology because they do not like the style of teaching. So perhaps the physical sciences could be made yet more attractive by greater attention to the method by which material is presented.

Turning to the perceived usefulness of the different science subjects, the teachers' replies indicate that they view physics and chemistry as being more useful for careers than biology. Probably they also think that parents endorse this ranking, since they believe that parental influence operates more frequently to persuade pupils to choose physical than biological science.

The opinions of the teachers that pupils rarely choose science subjects because of their relevance for future family life is very perturbing. Science can, and should be related to pupils' present and future everyday lives, but the indications are that school science is too far removed from everyday experiences. Biology, probably because it entails studying the human body, is judged to be less remote than the physical sciences.

The teachers in the sample believe that sometimes pupils' subject choices are influenced by the sex typing of a subject. For example, they believe that an important reason why boys choose physics is that physics is a boys' subject, and that girls drop physics for the same reason. Furthermore, the teachers perceive that boys and girls expect to take up very sex stereotyped occupations. The teachers believe that boys are likely to seek a career for which physics is useful, whilst girls are likely to have more need of biology for their future occupation. In addition, girls are more likely to view physics as being irrelevant to their future family life than are boys. This implies that the teachers believe that it will be the boys who change plugs, mend fuses, and do similar tasks when they marry.

9.4 TEACHER EXPECTATION AND TEACHER JUDGEMENT

The results presented in section 7.4.1.2 to 7.4.1.5 give clear indications that, within the sample of teachers investigated, written work attributed to a girl was often given lower grades than identical work attributed to a boy, and the female teachers often gave higher grades than the male teachers. Consequently, the combination of boy's work marked by a female teacher was most likely to produce a generous assessment, whereas girl's work marked by a male teacher was most likely to produce a severe assessment. The outcome of the latter combination of pupil sex and teacher sex gives grave cause for concern, as the majority of physical science teachers are men. Sex biased marking practices could be discouraging girls from studying science and disadvantaging those girls who do choose to study science.

9.4.1 Effect of pupil sex

9.4.1.1 Expectations and beliefs

In the marking exercise, a boy author was judged to be significantly more suitable for O level physical science courses than was an identical girl author. This biased assessment occurred regardless of whether the written work being evaluated was of above average, average, or below average quality. Boy authors who produced work of average and below average quality were also judged to be significantly more suitable for CSE physical science courses than were identical girl authors. Unfortunately, physical science CSE suitability proved to be an ambiguous variable when applied to the sample pair representing a high standard of work. O level suitability, together with CSE suitability, were included as direct indicators of teacher expectation. The results clearly signify that the teachers generally expressed higher expectations for the boy pupils than for the girl pupils.

9.4.1.2 Work and pupil characteristics

In the marking exercise, the work of boys was rated significantly higher than that of girls for scientific accuracy and understanding of principles in two of the three samples of experimental write-ups. Both these variables correlate highly with the overall mark awarded to a piece of work, which suggests that they are particularly central factors in assessment. Thus, the findings suggest that the teachers' judgement of factors central to the marking exercise was more likely to be influenced by the sex of the pupil than was their judgement of less relevant factors.

Only one experimental write-up, the one of average quality, was given a significantly higher mark out of 10 when written by a boy than when written by a girl. Although the distribution of marks awarded to the boy and to the girl should theoretically have overlapped exactly, in actuality 21% of the combined area covered by both distributions was nonoverlapping. If a similar degree of nonoverlap in the marks awarded to equivalent standard work from boys and girls occurred during marking of O level scripts, then 50% of the boys' scripts would receive higher marks than 62% of the girls' scripts. Such substantial differences in marks awarded would presumably result in work of similar standard being awarded different grades.

The teachers displayed very pronounced sex bias in their opinions concerning the pupil characteristics that they were asked to assess. For all three sample pairs, boys were rated significantly higher than girls on aptitude for science, attitude towards science, and interest in science. The different ratings awarded to girls and to boys for aptitude clearly indicate that the teachers regarded the boy pupils to be better suited for science studies than the girl pupils. The other two variables, attitude and interest, also provide an indirect indication of a pupil's potential for science, since a favourable attitude and interest in science is likely to result in greater motivation and greater strivings in both present and future assignments. Thus the teachers' sex biased ratings of aptitude, attitude and interest provide additional support for the view that they held higher expectations for the boy pupils than for the girl pupils.

Present judgements and future decisions concerning pupils' academic careers are likely to be based on perceptions such as those discussed above. Thus the lower scores awarded to girls for scientific accuracy, interest, O level suitability, etc., not only signal the sex biased marking practices and expectations of some science teachers, but also suggest that they will most likely continue these practices.

It could be argued that the teachers' higher ratings of boys' work and attributes merely reflects the superior performance and attitudes towards science customarily associated with boys. Certainly there is much evidence that, from about age 13 upwards, boys are more interested in science, better at science and more likely to continue studying science (Kelly, 1981). However, there is little evidence that the academic attainment of boys in science is noticeably greater than that of girls at the lower secondary level. A recent study by Smail and Kelly (1984) has shown that girls and boys enter secondary school with approximately equal knowledge of science. If this is the case, then the findings of the present study, that the work and attributes of a first year pupil are evaluated more highly when linked with a boy's name than when linked with a girl's name, must prompt us to question the magnitude, if not the very basis of boys' customary superior attainment in science. It is possible that the superior academic achievement of boys in science is largely maintained by the biased expectations, perceptions and judgements of science teachers. The results presented in section 7.4.1.2 suggest that even if there are differences between boys' and girls' aptitude for science, their interest in science, their attitudes towards science, etc., teachers may well be further magnifying these differences. The suggestion that the distorted perceptions of teachers can lead to differential expectations for boys and girls in science, differential behaviour towards boys and girls in the laboratory or classroom, and sex biased assessment of their work needs to be considered very seriously.

9.4.2 Explanations

9.4.2.1 Findings requiring explanations

Sex bias in the evaluation of male and female achievements has been recorded in a number of different contexts. For more than half a century, American researchers have been reporting a discrepancy between the marks awarded to boys and to girls when teachers set and mark exercises and examinations themselves (see section 2.5.2.1). When the marks of pupils of similar intelligence are compared, it is found that the girls generally receive higher marks than the boys. Even when the boys obtain better grades on objective aptitude tests, the girls still obtain better marks from their teachers in school examinations. This tendency for teachers to favour the work of girls has been detected at both the primary level and secondary level, and in a range of school subjects, e.g. reading, language, arithmetic.

The marks that teachers award to the work of boys and girls in marking experiments seem to differ from the marks that they award in a natural school setting. When several non-cognitive variables, including pupil sex, are investigated together in a marking experiment, their combined effect upon the teachers' marks is highly complex (see section 2.5.2.2.2). Pupil sex rarely produces a main effect. The interaction effects involving pupil sex that have been reported are confusing and sometimes contradictory. More research into the effect of combinations of non-cognitive variables is required before generalizations can be drawn regarding interaction effects that involve pupil sex.

When teachers are engaged in marking experiments that only manipulate the single variable of pupil sex, then the available evidence (see section 2.5.2.2.1 and the present findings) shows that work attributed to a boy is generally viewed more favourably than identical work attributed to a girl. Furthermore, the present research has shown that not only the work, but also a number of the personal characteristics of a boy author are rated more highly than those of an identical girl author. Although this investigation has shown very clearly that science teachers tend to overvalue the work and personal characteristics of boys, this simple finding is actually an overgeneralization that hides differing patterns of marking. For instance, the number and range of variables on which boys and girls received significantly different ratings varied according to the standard of the work being assessed. Work of average standard evoked a greater number of sex biased judgements than did work of above average or poor standard. Another variable that determined teachers' marking patterns was their qualifications for the task asked of them. Only teachers with appropriate teaching experience, i.e. they had taught chemistry and/or integrated science, displayed the sex biased marking patterns discussed above. Science teachers who were not well qualified to mark samples of chemistry work, produced very different patterns of biased marking. When the standard of the work was high, they favoured the work of a girl; when the standard was average, they displayed no sex bias at all; and only when the standard of the work was poor did they favour the work of a boy.

Goldberg (1968) and a number of other researchers have conducted experiments with very similar experimental designs to the marking exercise described in this thesis, the main difference being that samples other than pupils' written work were presented for evaluation. Also the subjects were usually college students rather than teachers. In spite of these differences, the findings were broadly similar to those of this study (see section 2.5.1.2). The achievements of men (i.e. journal articles, paintings) were frequently judged to be better than identical achievements from women. Moreover, a male author/artist was judged to be more competent than a female author/artist. The upgrading of articles by male authors tended to be most pronounced in masculine fields. Research has shown that a number of variables influence the way in which the achievements of men and women are evaluated, including the expertise of the rater, the sex appropriateness of the achievement, the level of the achievement, and any ambiguity concerning the qualifications or status of the person who produced the work.

9.4.2.2 Explanatory hypotheses

The findings discussed in the previous section all refer to different aspects of the same phenomenon - sex biased evaluations of male and female achievements. A satisfactory explanatory hypothesis should ideally be able to account for the various facets of this phenomenon. The literature suggests a number of hypotheses that have been or could be used to explain the differential evaluation of male and female achievements. Not surprisingly, the hypotheses vary in the emphasis they place upon differences between evaluators, differences between the males and females being evaluated, and differences between evaluation methods. In the following paragraphs a number of hypotheses are described and their relative strengths and weaknesses are discussed.

Four hypotheses have already been proposed in the literature review (section 2.5.2.1.1) to account for girls' greater success in exercises and examinations which are set and marked by their teachers. Since these hypotheses have been discussed with respect to one aspect of the differential evaluation phenomenon, their capacity to account for other aspects of the phenomenon should also be considered. As the four hypotheses have already been explained in detail, they will only be described briefly here.

The first hypothesis suggests that female teachers discriminate against boys, and since the majority of primary teachers are female, their biased treatment of boys could explain the significantly poorer marks received by boys at the primary level. At the secondary level, where a higher proportion of the teachers are male, the boys gradually improve their marks relative to girls. This particular hypothesis fails to explain satisfactorily the bias in favour of girls recorded in natural school settings, since the work of girls is assessed more favourably than that of boys by both female and male teachers. Thus it is not just female teachers who discriminate against boys, but both sexes. Turning to other aspects of the differential evaluation phenomenon, the explanation offered by this first hypothesis is totally unsuitable. It is illogical to explain the tendency of respondents to overrate the work of males in assessment experiments by arguing that female respondents discriminate against males' achievements. The shortcomings of this hypothesis are further emphasized by the fact that some researchers, e.g. Goldberg (1968), only used female raters in their assessment experiments and yet still recorded a bias in favour of men's articles.

The second hypothesis suggests that girls overachieve at their school work compared to boys of similar ability, and a third hypothesis suggests that teachers favouring of girls' work results from the operation of halo effects. When teachers assign marks to girls for their academic work, they are unduly influenced by girls' better attitude towards school work and their superior conduct in the classroom. A fourth explanation is based upon recent evidence that boys are better at multiple choice questions, whereas girls excel on essay type questions. This last explanation satisfactorily accounts for boys' higher grades on objective aptitude tests, and girls higher marks on non-objective type questions set by teachers. In contrast, the over-achievement by girls hypothesis totally fails to account for boys' superior performance on aptitude tests. However, the halo hypothesis does seem to offer a plausible explanation of boys' and girls' reported strengths and weaknesses on school based and aptitude tests. As with the first hypothesis, the three additional hypotheses just mentioned all fail to explain adequately the tendency for the work of a male to be favoured in assessment experiments. The higher ratings awarded to the work of a boy compared to that of a girl by teachers in a marking exercise cannot be explained by arguing that boys perform better on objective type questions, since the samples of pupil work used in the marking exercise were experimental write-ups and essay type answers. The overachievement hypothesis and the halo effect hypothesis could partly account for the lower marks awarded to girls in marking experiments, but the hypotheses do not adequately explain why the work of boys is positively favoured in marking experiments.

Besides the four hypotheses that were originally proposed to account for girls' better level of attainment in natural school settings and which have been discussed above, another three hypotheses are worthy of consideration. It could be argued that girls perform better at their school work because they mature more rapidly than boys. However, this hypothesis fails to explain why boys perform better on aptitude tests, and why fictional males are assessed more favourably than identical fictional females. Differences between assessors might form the basis of a second explanation. It could be argued that the way that assessors perceive, evaluate, and rate samples of males' and females' work is largely determined by the characteristics of the group of assessors, e.g. their age, their competence to rate the samples, even the sex ratio of the group. Reference to factors of this description could help to explain the inconsistent evaluation of male and female achievements that has been reported. However, such explanations would be severely restricted by the absence of any unifying theoretical stance, and would be increasingly difficult to apply to future findings. A third explanation might focus upon differences between the gender images of the subject areas being assessed. Using similar reasoning to that employed in section 1.2, it could be argued that assessors hold different expectations for males and for females depending upon the gender image of the work being assessed. Women are expected to produce better work on feminine topics, and men are expected to excel on masculine topics. Not only do the gender connotations of a subject area induce greater expectations for the corresponding sex, but the expectations are so powerful that they influence the assessors' perceptions and judgements. This leads to the work of one sex receiving more favourable ratings than that of the other sex. This hypothesis would satisfactorily explain boys' higher ratings in the marking exercise. Chemistry is a boys' subject, and so the teachers may well be inclined to give higher marks to samples of chemistry work that have apparently been produced by boys. The hypothesis would also explain some of the results reported from Goldberg-type experiments. However, the hypothesis cannot explain the indiscriminate marking up of girls' work in schools. This even occurs in subjects like maths, which most teachers regard as a neutral or slightly masculine subject.

9.4.2.3 Combining the hypotheses to account for research findings

The seven different hypotheses discussed above each attempt to explain, but with varying degrees of success, one particular aspect of the differential evaluation phenomenon. A few of the hypotheses go some way towards explaining a couple of the different aspects of the phenomenon, but not a single hypothesis satisfactorily accounts for all the different facets of the phenomenon being discussed in this thesis. It would appear that either the phenomenon is too complex and divaricate, or that the explanatory hypotheses are inappropriate or too simplistic. Assuming that the phenomenon is complex, then it would seem to be appropriate to develop a theory that combines the assumptions made and implications suggested by several of the simple hypotheses already discussed.

A theory that is formed by combining elements of the halo hypothesis and the sex stereotyped expectations hypothesis would appear to possess considerable explanatory potential. The arguments from the halo hypothesis could primarily explain the tendency of teachers to favour the work of girls in classroom settings, and the tendency of assessors to favour the work of males under experimental conditions could be primarily explained using arguments from the sex stereotyped expectations hypothesis. The choice of these two hypotheses was guided by the following considerations. Of the five hypotheses proposed to account for school teachers' favouring of girls, the female teacher bias hypothesis, the overachievement of girls hypothesis, and the physical maturation hypothesis cannot satisfactorily explain reported findings; and the type of question hypothesis cannot be applied to the marking experiments. Only the halo hypothesis appears to offer a satisfactory explanation of school findings, and be sufficiently adaptable to be relevant to other aspects of the differential evaluation phenomenon as well. The emphasis of the halo hypothesis upon factors that influence teachers' judgements is thought to be particularly apposite considering teachers' tendency to favour either of the two sexes under different circumstances. Of the various hypotheses under consideration, the sex stereotyped expectations hypothesis is judged to be the most satisfactory hypothesis to account for assessors favouring of males' work and personal characteristics under experimental conditions. The sex stereotyped expectations hypothesis and the halo hypothesis combine together particularly well since both involves teachers using preconceived beliefs and expectations to interpret information, with the result that their judgements are influenced by extraneous factors.

As mentioned earlier, girls' better performance at school work can be explained using the halo component of the composite theory. Teachers perceive that girls conform better than boys to classroom rules and norms. They believe that girls are more conscientious over their work, and they expect the work of a girl to be more thorough and neater than that of a boy (see section 7.2.1.2). When teachers come to mark the work of their pupils they match each piece of work with the pupil who produced it, with the result that the mark awarded tends to reflect more than just the work's quality. The work of pupils who are viewed favourably, perhaps because they are well behaved, conscientious, hard working, is also viewed favourably. The work of pupils who are problematical will tend to be viewed unfavourably. Because teachers tend to stereotype the behaviour and work characteristics of boys and girls, the pupils that teachers view favourably will mostly be girls, and the pupils who are viewed as less cooperative will tend to be the boys. When the pupils sit an aptitude test, the marking is entirely objective and unbiased. Halo effects cannot operate and so the work of girls is not marked up. Consequently the marks of boys and girls are often very similar. On those occasions when boys achieve the better marks, this could be due either to the boys being more able or to their greater competence at objective tests.

Halo effects appear not to operate when appropriately experienced teachers mark the work of fictional pupils. The experimental finding that the work of boys is generally favoured over that of girls can best be explained using the sex stereotyped expectations component of the composite theory. It can be argued that since chemistry is a boys' subject, the teachers expect boys to produce better work than girls. Their expectations influence their perceptions and judgements of the work samples that they assess, with the result that the work of boys tends to be marked up and that of girls to be marked down. It is noteworthy that the bias displayed by the teachers was most extensive on the work sample pair of average quality and least on the below average work. Presumably, the teachers unconsciously acknowledged that the boy who produced the poor quality work was unlikely to progress very far in science related fields. There was thus little incentive to favour his work compared to that of the girl. The bias associated with the good quality work is particularly interesting, since it was focused upon the pupil characteristics rather than the work variables. It would appear that teachers accept that there are some bright girls who can produce work of a high quality, yet they find it hard to accept that such girls are as interested and committed to science as boys. Even thought they recognise the achievements of bright girls, they still cling to the belief that boys are better suited to science studies.

Sex stereotyped expectations can also help to explain a number of the findings associated with Goldberg-type experiments. Some researchers have reported that work attributed to a woman tends to be favoured when the work refers to a feminine sphere, whilst work in a masculine or neutral field tends to be rated higher when it is attributed to a man. Such findings are consistent with the sex stereotyped expectations component of the composite theory.

To explain the marking patterns produced by teachers without appropriate experience for the task asked of them, the two elements of the composite theory have to be combined in varying proportions. It is assumed that the inappropriately experienced teachers did not possess sufficient knowledge of the topic presented for assessment to be able to accurately mark the work, but that they could assess the overall standard of the work. Since their marking was inevitably very subjective, they were heavily dependent upon extraneous factors. On being presented with the work sample pair of above average quality, the teachers probably recognised that high standard of the work. When such work was associated with a girl's name, they judged that the author was academically able and assumed that she was conscientious. They would also expect her work to be neat, her prose to be good and her spelling to be accurate. Such expectations were probably confirmed by the work sample. To compensate for their lack of expertise in the subject area, the teachers probably payed undue attention to extraneous factors such as perceived neatness, when awarding marks for the work related variables. This tactic would tend to lead to higher ratings being given to the work related variables when the sample was attributed to a girl. Since the pupil characteristics, with the exception of aptitude and O level suitability, could be assessed more accurately, their ratings showed less bias than the ratings of the work characteristics, In fact, the inappropriately experienced teachers' ratings of the pupil's attitude and interest were less biased than those of the appropriately experienced teachers. When inappropriately experienced teachers were confronted with the average quality work, since the standard of the work was obviously not so high, the halo effects described above did not operate so effectively to advantage the work of a girl. The overall result was that no bias was detected. When the inappropriately experienced teachers marked the poor quality work a different set of expectations were in force. The teachers found it acceptable for a girl to be poor at science, but they found it unacceptable for a boy to be poor at science and so they marked up the work of a boy.

The arguments presented above to explain the marking patterns of the inappropriately experienced teachers suggest that two sets of expectations and influences were in operation. Sex stereotyped expectations arising from the belief that chemistry is a boys' subject dictated that the work of boys should be marked up. Sex stereotyped expectations that girls produce neat, thorough work, reinforced by corresponding halo effects, dictated that the work of girls should be marked up. When inappropriately experienced teachers were marking, the two conflicting sets of expectations reached a balance at different points depending upon the standard of the work. In contrast, the apporpriately experienced teachers were cognizant of the work that they were marking and thus halo effects were not resorted to. Only the expectations that favour the work of boys operated.

The question of whether science teachers mark the work of boys and girls differently in their day to day teaching is difficult to assess from the results reported in this thesis. Theoretical considerations would seem to suggest that girls should be marked more favourably than they were in the experiment. However, whether any advantage to girls that arose from halo effects would be sufficient to balance the experimentally demonstrated bias towards boys is difficult to predict. Since teachers generally seem to hold higher expectations for boys, and since they firmly believe that boys' attributes are better suited to science studies than are those of girls, then it is likely that the bias in favour of boys does operate under natural conditions, but more research is needed to ascertain whether this supposition is correct.

The two experimental findings that the appropriately experienced teachers showed most bias when marking the average quality work, whilst the inappropriately experienced teachers showed no bias when marking the same work are particularly worrying. They indicate that sex stereotyped expectations exert most influence when work of average standard is being marked, but that halo effects exert little influence. This suggests that bias in favour of boys is very likely to operate under natural conditions when the work of average pupils is being assessed. Support for this conclusion is provided by studies of girls who enter engineering. Girls of average ability who eventually become technicians encounter much more discouragement and discrimination than do brighter girls who take engineering degrees (Chivers & Marshall, 1983; Newton, 1983). The present work suggests that perhaps girls who study science fare similarly. If this is the case, the consequences deserve further consideration and investigation.

9.4.3 Educational implications

Assessment in schools serves a variety of purposes (see Matys, 1970). At the individual classroom level, assessment can be used to help the teacher to teach better and the pupil to learn better. Marks can benefit pupils by providing them with feedback about their academic progress in a subject, and by motivating them to work harder. In addition, teachers gain information about individual pupils' problems and weaknesses. Marks also enable teachers to group pupils for instructional purposes, and to provide guidance on further education or employment opportunities. However, unless assessment is impartial, responses and decisions will be based upon false information. The consequences for an individual pupil's educational and occupational career could be quite serious.

The marks that a pupil obtains for homework or classwork can represent a significant proportion of the personal feedback that the pupil receives from a teacher. This is especially true if the class is large and/or if the teacher only sees the class for a couple of lessons each week, which could be the case in a secondary school that splits the teaching of science in the lower school between the speciality science teachers. Marks not only convey information to a pupil about the teacher's assessment of the value of the pupil's performance, but the pupil can also interpret them as an indication of the teacher's evaluation of his/her value or worth (Sharp & Green, 1975). Thus good marks can enhance self-evaluation, self-esteem (Craparo et al., 1981), and expectations for future success. Pupil expectations are of considerable importance since they are an important variable in determining subsequent achievement (Smead & Chase, 1981).

Favourable feedback can have positive effects upon children's performance (Feather, 1966; Tait et al., 1973). In contrast, low marks can discourage pupils, they can lose interest in a subject and cease to work at it. Girls appear to be more discouraged by low marks than boys. Edwards & Wilson (1958) showed that girls were more likely to cease studying chemistry and physics as a consequence of poor marks than were boys. This is particularly worrying since physics and chemistry are notoriously difficult subjects (Nuttall et al., 1974). If physics was marked more leniently, Bridgham (1973) suggests that the number of girls choosing to study the subject in the United States, would increase by 80%. The effect upon boys would be negligible. This finding clearly suggests that girls are more discouraged and deterred by the perceived difficulty of a subject than are boys. The same conclusion was also reached by Keys and Ormerod (1977). They found that girls' preferences for subjects were more closely related than those of boys to their perceptions of the difficulty of the subjects.

Because girls often get higher marks than boys in the arts subjects, the contrast between the marks of arts and science subjects is likely to be greater for girls than for boys. Thus girls are more likely to perceive the physical sciences as difficult subjects. Any tendency on the part of science teachers to mark girls' work more severely than that of boys would disproportionately deter girls from science and further widen the gap between boys' and girls' uptake of science.

Girls have a poorer academic self-image than do boys (Barker Lunn, 1972; Sears & Feldman, 1974), and are less confident about their academic ability (Oetzel, 1967). Recent British work has generally confirmed this lack of self-confidence displayed by girls (J. Best, personal communication, 1981; DES, 1980). Sex differences in self-confidence appear to be dependent upon the ability areas investigated. An American study has shown that women have lower self-confidence than men on spatial-mechanical and creativity tests (Lenney, 1981). Females' lack of self-confidence in academic spheres could explain why low marks discourage girls more than boys. Peterson et al. (1980) report that pupils' self-concepts in science are interrelated with their success in science as measured by grades. Whether self-concept was the precursor or result of success was not investigated. But if girls have little self-confidence in science and they are given low marks, neither causal linkage could improve girls' position in science.

Dweck & Bush (1976) found that failure feedback from an adult impaired girls' experimental task performance and persistence on a subsequent task, whereas Scarbo (1979) found that, in an experimental situation, adult praise led girls to solve problems faster and gain higher scores on a related substitution task. These two examples demonstrate the motivational value of praise and social approval to girls. In the classroom, girls seek social recognition in the form of high marks (Edwards & Wilson, 1958). Social approval improves girls' academic performance, whereas boys improve more under competitive conditions (Mischel, 1967). Because social approval is such an effective reward for girls, the motivational value of good marks is considerable. Thus when girls receive low marks they are not only being denied social approval, but they are also being deprived of an important source of achievement motivation.

Level of school performance can provide additional motivation by contributing towards interest and enthusiasm in a subject, as well as improving attitudes towards it. Many studies have been reported in which pupil attitudes and interest were significantly related to academic success. The causal direction of this relationship has been investigated by Eisenhardt (1977). Measures of interest and achievement scores in four academic areas (science, maths, social science and English) were collected from over 70,000 pupils over a two year period.

> The findings suggest that the predominant causal sequence is from changes in achievement levels causing changes in interest levels across all sample groups in each of the four academic areas ... more often than changes in interests cause changes in achievement.

This study indicates that low marks not only present a record of poor cognitive attainment, but they will also depress affective responses towards the subject. Once achievement and attitudes are poor, they will most likely reinforce each other through feedback loops.

In addition to using marks to reward pupils for past performance and to motivate pupils to perform well in the future, teachers also form expectancies for pupils' future performances. There is thus a danger that any initial bias within a teacher's marking gets accentuated over time. Furthermore, because pupil ability and performance are such important factors, they tend to influence perceptions of many other attributes as well (Nash, 1973; Solomon & Kendall, 1977).

In their capacity as arbitrators of what is success and what is failure in a particular school subject, teachers wield considerable power and influence over each pupil's prospects in that subject. No doubt teachers believe that their evaluations of pupils are fair and unbiased. However, many studies have provided indications that such is not always the case. This study alerts science teachers to the possible biasing influence of pupil sex.

CHAPTER 10

CONCLUSIONS AND IMPLICATIONS FOR SCIENCE EDUCATION

10.0 Contents

10.1 INTRODUCTION

In the past, too little attention has been paid to the attitudes and actions of the science teacher. Most studies concerned with girls' under-representation and under-achievement in science have focused upon differences between boys and girls. But pupils do not learn science in isolation; they are generally taught by teachers. Thus the teacher's influence on the learning process should also be studied. If teachers' attitudes and expectations predispose them to treat boys and girls differently, it would be surprising if boys and girls were not to respond differently.

This study set out to determine (a) whether science teachers hold different expectations for their male and female pupils, (b) whether their expectations are linked with sex stereotyped beliefs about pupil and subject characteristics, and (c) whether sex differentiated expectations are associated with sex biased assessments of pupils' work and personal attributes. Attention was also directed to identifying personal and educational variables that might be indicative of a teacher's propensity to differentiate between pupils on the basis of their sex. Finally, it was hoped that the results from the various investigations would throw more light upon the role that teachers might inadvertently play in depressing girls' level of attainment in science and discouraging them from continuing science studies.

10.2 SUMMARY OF FINDINGS

The same five topics of investigation that were identified in Chapter 3, and which have helped to structure the whole thesis, are again used in this section.

10.2.1 Sex typing of science

Physics, chemistry, biology and maths form a closely related group of school subjects that can broadly be termed 'science' subjects.

Teachers' ratings on semantic differential scales indicated that they regard this group of subjects to be more scientific, logical, factual, routine, complex and important than a range of other secondary school subjects. The 'science' subjects are also generally judged to be more numerical and more masculine than other school subjects (section 7.1.1.1).

Science teachers' responses regarding the gender connotations of the science subjects depended upon the type of measuring scale used. The use of a simple, explicit 3-point scale (masculine-neutral-feminine) tended to produced neutral ratings (section 7.1.4.2). More differentiated responses were obtained by using gender rating scales with a greater number of rating positions (section 7.1.1.1). The use of indices that measured gender connotations in a more oblique manner produced the clearest evidence that science teachers do sex type the science subjects (section 7.1.2).

Secondary science teachers' responses to discriminating gender scales indicated that they regard physical science subjects, i.e. physics and chemistry, to be masculine subjects. Physics was judged to be slightly more masculine than chemistry. Biology was judged to be very slightly feminine. Thus, the three common science subjects can be ranked in order of masculinity - physics, chemistry, biology - with physics being the most masculine subject (section 7.1.2).

The masculine image of physics is linked in science teachers' minds with a number of other characteristics that the subject is believed to possess. Elementary linkage analysis produced a clearly defined 'masculine' cluster that contained five characteristics: masculine, technical, mechanical, mathematical, and concern with objects. These same five characteristics also appeared in the gender cluster for biology, which suggests that similar associations probably underlie the gender image of all the science subjects (section 7.1.3.2).

Science teachers' views about the gender connotations of the three science subjects were not significantly different to the views of other secondary teachers (section 7.1.1.2). However, secondary teachers (including science teachers) viewed physics to be significantly more masculine than did primary teachers. They judged biology to be significantly more feminine than did primary teachers. Thus secondary teachers were more inclined than primary teachers to sex type the science subjects (section 7.1.1.3).

Science teachers expressed the opinion that the factor that is primarily responsible for giving the physical science subjects a masculine image is the number of male scientists. Not only are most scientists men, but in addition scientists are usually protrayed as men by the media, e.g. adverts, films, comics. Tradition, stereotyping and social pressures were also perceived to be important factors contributing to science's masculine image (section 7.1.4.1).

Science teachers picture a physicist to be a very different type of person from a biologist. The sample teachers indicated on semantic differential rating scales that a physicist is quite probably male, whereas a biologist is equally likely to be male or female. They also judged that a physicist is significantly more likely than a biologist to be good at maths, logical, objective and competitive; and significantly less likely to be sociable, emotional and humanitarian. These findings suggest that science teachers mainly associate physical scientists with stereotypically masculine qualities (section 7.1.5).

10.2.2 Sex Stereotyping

10.2.2.1 Written Work of Girls and Boys

Most teachers recognise differences between the written work of girls and boys. When asked to list features that they consider to be typical of the written work of each sex, 85% of a sample of science teachers responded. Furthermore, nearly three quarters of the sample claimed that they can generally distinguish between the written work of girls and boys. A significantly higher proportion of male science teachers (80%) than female science teachers (57%) indicated that they believe they can identify the work of girls and boys (section 7.2.1.1).

The features mentioned by science teachers who listed differences between the written work of boys and girls fell into the following categories.

1. Over 90% of the teachers noted differences in the appearance of girls' and boys' written work. There was unanimous agreement that girls' work tends to be neat and well presented, with good diagrams. In contrast, boys' work was described as being untidy and poorly presented.

2. Over 30% of the teachers perceived differences between the sexes in their approach to their work. Girls were seen as being conscientious and boys as being careless.

3. Aspects of handwriting were mentioned by over 20% of the science teachers. They judged girls' writing to be rounded, neat and easily legible.

4. Nearly 50% of the teachers commented upon the quality of the work produced by each sex. Boys were credited as displaying more understanding than girls, and the teachers were agreed that boys produce accurate work. There was disagreement over the accuracy of girls' work, but not over its thoroughness.

5. About 30% of the teachers noted the amount of work produced by boys and girls. All comments referred to the length of girls' work and the brevity of boys' work (section 7.2.1.2).

A group of non-science teachers referred to very similar features when describing the written work of boys and girls. However, the non-science teachers were less inclined than the science teachers to mention differences in the quantity of work produced by boys and girls, and differences in the quality of its content (section 7.2.1.2).

The findings recorded above suggest that either substantial differences do exist between the work of boys and girls, or more likely that popular cultural stereotypes concerning the characteristics of boys and girls and their work are accepted by most teachers.

10.2.2.2 Preference for Subject Characteristics

Teachers' perceptions of the type of subject that appeals to boys and to girls were investigated using bipolar semantic differential rating scales. It was found that science teachers believe that boys and girls prefer significantly different subject characteristics. Most noticeably, boys were thought to prefer subjects than can be described as numerical, science, logical and masculine; whereas girls were thought to prefer subjects that are verbal, arts, intuitive and feminine. In addition, practical and factual subjects were considered to be more attractive to boys, whilst simple and creative subjects are more attractive to girls (section 7.2.2.1).

Factor analysis of the replies suggests that science teachers link the two variables 'arts' and 'feminine' together closely when describing subject characteristics preferred by girls, but do not make a similar linkage between the variables 'science' and 'masculine' when describing subject characteristics preferred by boys. A correlation scatter plot corroborated this finding (section 7.2.2.4).

Secondary school teachers (including science teachers) tended to express more sex stereotyped views about pupils' preferences for subject characteristics than did primary and middle school teachers. For example, secondary teachers were more committed to the view that girls prefer subjects that are simple and verbal. Their beliefs about the attraction of a subject's gender image were even more stereotyped. Secondary teachers were not only more extreme in their view that girls prefer feminine subjects, but also that boys prefer masculine subjects (section 7.2.2.2).

10.2.2.3 Females' Social Roles

Science teachers' views of adult sex roles were investigated by administering a short-form Attitudes to Females' Social Roles questionnaire. 40% of the teachers obtained low scores (half the maximum score or less), indicating that they hold traditional, sex stereotyped attitudes towards women's role in society. However, the fact that nearly a quarter of the teachers, especially women, expressed very liberal attitudes should not be overlooked.

Science teachers' responses to the individual items of the questionnaire established that approximately one quarter of them agreed that:

1. Women's careers and earning capacity are secondary to those of men.
2. The man is the dominant and controlling partner in a marriage.

In both of the pilot studies and the main study, women obtained significantly more liberal scores than did men. The results from the second pilot study and the main study also showed that teachers under 40 gained significantly more liberal scores than did teachers who were 40 and over (section 7.2.3.2).

10.2.2.4 Importance of Subjects

To investigate the perceived value of different subjects for boys and for girls, teachers were asked to rate the importance of qualifications in a range of optional subject areas to pupils' future lives. Secondary teachers did not regard all school subjects to be of equal importance for boys and girls. They indicated that science and technical subjects are of greater importance to boys. Home economics and commercial/business studies were judged to be of greater value to girls (section 7.2.4.1). In addition, science teachers also rated humanities and creative arts to be of greater value to girls than to boys.

Science teachers' views about the importance of different subject areas were more sex differentiated than were those of teachers of other subjects. Not only did they regard the acquistion of qualifications in the science subjects to be significantly less important for girls than for boys, but they even judged science to be of significantly less value to girls than did teachers of other subjects (section 7.2.4.2).

10.2.3 Attribution Patterns

10.2.3.1 Causes of Success/Failure at Science

The hypotheses that science teachers mostly attribute boys' success at science to stable internal factors, e.g. ability, but that they mainly attribute girls' success to unstable factors, e.g. effort or luck, were not confirmed. With regard to their principal teaching subject, teachers were asked to rate the importance of a number of factors to girls' success or to boys' success. Their combined replies showed that science teachers believe that internal factors, such as a pupil's effort, ability and interest in the subject, contribute more to success in science than do external factors, such as family support, out-of-class experience. Similar attribution patterns were used to explain the success of both boys and girls in each of the three main science subjects, with one exception. Pupil effort was judged to contribute significantly more to girls' success in physics than to boys' success (section 7.3.1).

The hypotheses that science teachers mostly attribute boys' failure at science to unstable factors, e.g. lack of effort or bad luck, but that they mainly attribute girls' failure to stable internal factors, e.g. lack of ability, were not confirmed either. Teachers' replies to a scale that investigated the perceived importance of various factors that contribute to failure at science showed that internal factors, such as lack of effort, poor attitude and lack of interest in the subject, are believed to contribute most to failure in science. Similar attribution patterns were used to explain the failure of both boys and girls in each of the three main science subjects, except that subject difficulty was believed to contribute more to girls' failure in physics, and lack of relevant out-of-class experience to boys failure in biology (section 7.3.2).

10.2.3.2 Reasons for Choosing/Dropping Science

Teachers' perceptions of the reasons why boys and girls choose either to continue or to drop science subjects were investigated by asking them to rate how frequently they believe each of a number of reasons apply in their principal teaching subject. Their replies were combined to produce a composite picture.

Science teachers believe that boys and girls tend to choose science subjects for very similar reasons. A pupil's choice of science is thought to result from a mixture of school-based factors (e.g. liking of the subject content, liking of the teacher), and out-of-school pressures (e.g. the subject's relevancy for future careers, parental influence). However, tradition and the masculine connotations of science are believed to influence boys more frequently than girls in their choice of science. These two variables are thought not to be equally influential in persuading boys to choose the different science subjects. Science teachers believe that they most frequently operate to determine boys' choice of physics (section 7.3.3).

Science teachers believe that the most important reasons causing pupils to drop science are the same for both girls and boys. Pupils most frequently drop a science subject because they find it difficult or because they dislike it. Sex differences are thought to occur in the operation of some of the less important causes. Teachers think that boys are more likely than girls to drop science because they do not like the teacher, and girls are more likely to drop science because of its irrelevancy to their future family life. The influence of some factors is determined by both a pupil's sex and the science subject under consideration. Thus teachers believe that girls more often than boys drop physics and chemistry, whilst boys more often drop biology, because the subject is judged to be irrelevant for their future career, the subject is associated with the opposite sex, and because traditionally few pupils of their sex have studied the subject (section 7.3.4).

10.2.4 Teacher expectation and teacher judgement

To investigate whether pupil sex affects teachers' expectations and judgements, science teachers were invited to participate in a marking exercise. They were supplied with samples of pupils' work and asked to rate the work samples and their authors on a number of variables. However, pupil sex was varied so that the same piece of work was presented to half of the teachers as being the work of a girl and to the remaining teachers as being the work of a boy. From the ratings received from science teachers with appropriate teaching experience for the marking task, the following findings emerged.

1. Work attributed to a boy was generally rated higher for scientific accuracy and understanding of principles than was work attributed to a girl (section 7.4.1.2).

2. The teachers expressed higher expectations for boy authors than for girl authors. Boy authors were judged to be more suited for undertaking further courses in physical science subjects, specifically CSE and O level physical science courses (section 7.4.1.2).

3. Boy authors were judged to have significantly more aptitude for science than were girl authors (section 7.4.1.2).

4. Boy authors were judged to have significantly more favourable attitudes towards science and significantly greater interest in science than were girl authors (section 7.4.1.2).

4\. Female teachers generally gave higher ratings than male teachers (section 7.4.1.3).

5\. Teacher sex and pupil sex acted together in an additive manner to determine marks awarded. The combination of boy's work marked by a female teacher most frequently produced a generous assessment, whereas girl's work marked by a male teacher most frequently produced a severe assessment (section 7.4.1.5).

10.2.5 Relationships between variables

Investigations into relationships between the four topics reviewed above were inadvertently hampered by complex experimental designs and measuring scales. Only two relationships were identified, both of which involved Masculinity Index measures of sex typing. Masculinity Indices tended to interrelate with both sex stereotyping measures (boys' and girls' perceived preferences for science/arts and masculine/feminine subjects) and sex influenced marking patterns. The latter relationship, although only marginally statistically significant, did accord with theoretical predictions (sections 8.3.2.3 and 8.3.2.5).

Investigations into the effect of various independent variables upon the four main dependent variables were more fruitful.

1. Science teachers' perceptions of the masculine image of science subjects, their perceptions of a range of the characteristics possessed by science subjects, and their perceptions of scientists were influenced to varying degrees by their principal teaching subject and their sex. Most strikingly, physics and chemistry were viewed as being more masculine subjects by male teachers than by female teachers. Also physics teachers perceived physics to be a less masculine subject than did chemistry and biology teachers (section 8.1.2)
2. Science teachers' ideas about pupils' preferences for subject characteristics and about the importance of subjects for boys and girls, together with their scores on the Females' Social Roles scale, clearly indicate that female teachers hold less sex stereotyped views than do male teachers (sections 7.2.3.2 and 8.1.3).
3. Two variables were identified that relate with a propensity to mark boys and girls inequitably in the marking exercise. They are a teacher's social class background and the sex composition of the teacher's current school. Teachers from working-class backgrounds favoured the work of boys over that of girls to a significantly greater extent than did teachers from middle-class backgounds. Teachers who were currently

teaching in single sex schools favoured the work of boys over that of girls to a significantly less degree than did teachers who were teaching in coeducational schools (section 8.2).

10.3 THEORETICAL IMPLICATIONS

The structuring or restructuring of educational and/or psychological theory was not one of the principal aims of this investigation. The main intention was to elucidate and describe certain phenomena that were believed to bear upon the 'Girls and Science' problem. The ensuing findings related to this aim have been presented in Chapters 7 and 8. These findings do largely confirm the hypotheses proposed in Chapter 3. They also justify the attention that has been paid to the main topics that have formed the basis of this thesis.

In Chapter 1, Figure 1.1, a number of links between the separate topics were postulated. The existence of these links has not been firmly established in this investigation. Slight evidence in support of only two of the links is presented in this thesis. They are the sex typing of science - teacher expectation/judgement link (section 8.3.2.5), and the teacher expectation - teacher judgement link (section 8.3.2.2). The failure of this investigation to clearly demonstrate all the links was more likely due to the use of inadequate measuring scales and statistical analyses than to the non-existence of these links. In the light of the ample confirmation of the separate topics (with the exception of the attribution patterns), the links between them now appear even more plausible than at the start of the investigation. However, the failure of this study to confirm the theoretical framework drawn in Figure 1.1 in its entirety must be acknowledged. The separate topics have been spotlighted, but the causal links remain in darkness.

The marking exercise clearly demonstrated that science teachers do hold higher expectations for boys than for girls, and they do tend to favour the work of boys. However, the question of whether sex biased expectations are the main cause of sex biased judgements, or just one of many, has not been answered. Neither has the question of whether the effect of sex biased expectations upon sex biased judgements remains constant under different circumstances. Evidence from the marking exercise, and the findings of other researchers, indicates that such is probably not the case.

To explain the marking patterns obtained from appropriately qualified and inappropriately qualified science teachers for work samples of different standards, a composite theory has been proposed. This theory is formed from a sex stereotyped expectations hypothesis and a halo hypothesis. The two elements combine in varying proportions depending upon the context of the assessment and the qualifications of the assessor. In the marking exercise, the experimental findings obtained from the appropriately qualified teachers can best be explained using the sex stereotyped expectations component of the composite theory. It can be argued that since chemistry is a boys' subject, the teachers expected boys to produce better work than girls. Their expectations influenced their perceptions and judgements of the work samples that they assessed, with the result that the work of boys tended to be marked up and that of girls to be marked down. To explain the experimental findings obtained from inappropriately qualified teachers, both elements of the composite theory have to be used.

The two experimental findings that the appropriately qualified teachers showed most bias when marking the average quality work, whilst the inappropriately qualified teachers showed no bias when marking the same work are particularly worrying. They indicate that sex stereotyped expectations exert most influence when work of average standard is being marked, but that halo effects exert little influence. This suggests that bias in favour of boys is very likely to operate under natural conditions when the work of average pupils is being assessed. Whether bias also operates in favour of boys' work of above and below average quality under natural conditions is more difficult to conjecture.

Finally, the study has cast considerable doubt upon certain tenets of attribution theory. Contrary to the theoretical typology proposed by Weiner et al. (1971), luck was not one of the important factors used by science teachers to explain the success and failure of pupils in science. In the exploratory interviews none of the teachers mentioned luck as a factor that contributes to academic success or failure. In the first pilot, luck emerged as the least important of the factors under consideration. The relative unimportance of luck in teachers' attributions has also been found in other studies (Bar-Tal & Guttmann, 1981; Lorenz, 1982). If luck is in fact not a central and important attribution factor, then the attention and significance that has been paid to sex differences in the use of luck as a causal attribution can also be questioned.

Recent work has questioned the existence, pervasiveness, magnitude or interpretation of sex differences in self-attributions (McHugh et al., 1982). This study questions the pervasiveness of sex differences in attributions for others. Four hypotheses, that were highly feasible according to attribution theory, were proposed at the start of this investigation, but they were not confirmed. The obvious lack of support for these hypotheses could have arisen because the teachers were asked to consider the success and failure of pupils in general, rather than of a specific pupil. However, it is probably more realistic to view the nearly total lack of any sex differences in this study as yet another indicator that sex differences in causal attributions may not be as robust and clearly defined as has hitherto been accepted (see also section 2.3.1).

10.4 LIMITATIONS AND WEAKNESSES OF THE STUDY

Most of the points mentioned in this section have been discussed at length in previous chapters. Hence this section will merely note the major limitations and weaknesses inherent in the study. The reader is referred back to the original discussion of each limitation/weakness for full details about (a) its exact nature and extent, (b) an assessment of its theoretical consequences, and probable practical consequences, (c) the precautions or actions taken to counteract or combat the limitation/weakness, (d) indications that the limitation/weakness is not likely to crucially affect the findings.

Probably the most serious weakness of the study was the sampling procedure adopted. The original intention had been to use a single-stage cluster sampling technique, but the research actually relied upon opportunity sampling (section 4.4.1). Neither the selection of schools, nor the selection of teachers within schools, was totally determined by the researcher. Not all schools that were asked to help with the research agreed to do so; not all the teachers within a school or science department completed questionnaires, although sometimes they did (section 4.4.3.1). Although the return rate from some schools was low, there was no reason to suppose that a teacher's decision to complete a questionnaire or not was related to any of the variables under investigation.

The use of non-random samples can result in biased findings. Moreover, the data from non-random samples should strictly not be analysed using inferential statistical techniques. In this study, since it was easy to show that many of the sample characteristics investigated did not differ significantly from those of the whole population, the samples were judged to be representative of the population. Therefore the samples were treated as if they were random samples, i.e. inferential statistical methods were used.

Although great efforts were made to ensure that samples were sufficiently large, and generally they were, occasionally some of the sub-groups, e.g. female physics teachers, were under-represented and therefore possibly unrepresentative. With a small sample the sampling error is likely to be large, thus the statistics calculated from small samples are less accurate and stable than those calculated from large samples (section 4.4.1).

The unsatisfactory sampling technique used in this research threatened the validity of the findings. The differential selection of respondents was considered to constitute the main threat to internal validity. Internal validity was further threatened by the decision to occasionally compare results from non-identical semantic differential scales (Appendix 7.4). External validity was also threatened by selection bias, since the respondents' responses might have been unrepresentative of the population as a whole. However, reactive effects of data collection arrangements were seen to be the main threat to external validity. Comments received from the respondents gave no indications that they had ideas about the investigations that would invalidate the results. Although both internal and external validity of this study could have been better, it is argued that neither was seriously threatened (section 4.7). Thus it is permissible to generalize the findings of this study to other groups of science teachers, provided that the conditions are similar to those existing in this study. This stipulation, if strictly adhered to, does restrict the generalizability of the marking exercise findings to situations where the type of marking and the context of the marking is comparable to that of this study.

Determining the validity of the measurements proved to be very problematical (section 6.4). It was not possible to assess the concurrent validity of any of the scales used, and predictive validity was not applicable. Therefore attention was directed to the content validity and construct validity of scales. Unfortunately these validities are more difficult to assess.

The use of PGCE students to estimate the reliability of the measurements, a procedure which could be queried, failed to provide samples of adequate size (sections 4.4.3.2 and 6.3). The small sample sizes and the homogeneity of the respondents reduced the likelihood of obtaining high reliability coefficients. It should be noted that reliability coefficients were only obtained for those scales that appeared in the final form of the three main questionnaires.

Some of the decisions concerning data analysis introduced certain risks and weaknesses into the study. The use of parametric statistical tests involved making a number of assumptions about the data that were not tested (section 4.6.1). The setting of the acceptable significance level at 5% precludes the indiscriminate use of the findings from this research for building theories or advocating educational change. A higher level of significance is normally required before such conclusions are drawn. In this study, this requirement can be met, since those findings that are significant at the 1% or 0.1% level can be identified (section 4.6.2). The decision to analyse much of the data using multiple t tests was made with many reservations, since this approach can lead to difficulties in interpreting the results (sections 4.6.2 and 7.4.1.3). Awareness of this problem meant that considerable caution and conservatism had to be exercised in the interpretation and extrapolation of results. Finally, the statistical analyses performed to detect relationships between the dependent variables were singularly unproductive and hence possibly inadequate (section 8.3.4).

10.5 FUTURE RESEARCH

The research described in this thesis has answered a number of the questions that were initially posed. Perhaps inevitably many of these answers prompt further questions. Some of these questions, together with topics that were raised or hinted at in Chapters 1 and 3 but which were either not investigated or not satisfactorily investigated, will be considered in this section.

The most glaring gap in this research was the failure to demonstrate the links between the separate topics investigated. Not until these links have been described shall we have a complete understanding of the way that sex and gender influence the beliefs, perceptions and responses of science teachers. It is particularly important that the effects of beliefs and expectations upon teachers' responses be determined, and that the effects of teachers' responses, i.e. their judgements and behaviours, upon pupil attainment be investigated. Only when the complete chain of effects from teachers' beliefs, through their expectations and responses, to pupil attainment (see Figure 1.1) has been exhaustively researched shall we be fully aware of the many different ways that teacher effects influence both girls' perceptions and reaction to science, and also their achievements in science. We would then be in a better position to assess the extent to which teachers contribute to girls' under-achievement in science. A better understanding of teacher effects would also enable us to identify more precisely those beliefs and behaviours that are most deleterious to girls advancement in science. Once equipped with that knowledge, efforts could then be focused upon changing and improving those specific areas of teachers' attitudes and behaviour. In the meantime, we have little choice but to attempt a broad attack (see section 10.6) and hope that we touch the sensitive areas.

The desire to improve girls' experience of science by modifying teachers' views of girls in relation to science, and their treatment of girls in science classes, raises some particularly pressing problems. We may have many ideas about the ways that we would like teachers to respond to and treat girls in science, but we have few ideas as to how to achieve these ideals. Research is urgently required to produce effective techniques that sensitize teachers to issues of sexual equality, and help them to avoid inequalities in their classes. It is particularly important that these techniques should take into account the concern that girls of average ability are probably most vulnerable to biasing effects.

The findings from most survey and experimental research into sex stereotyping and sex bias are of considerable theoretical value, and they also direct our attention to specific mechanisms that are probably influential in real life situations. However, the extrapolation of experimental findings to classroom settings is always problematical. This study does not constitute an exception. Although sex biased marking patterns have been detected under experimental conditions, the question of whether such marking patterns also occur in natural settings has not yet been answered. The question requires urgent attention. A careful study of teachers' marks, linked with an investigation into teachers' verbal assessments and expectations of each pupil, should produce strong indications, if not firm answers. Evidence is also required concerning the applicability and relevance of other findings of this study to classroom practice.

Several questions arise from the finding of this study that science teachers tend to hold sex stereotyped views. Do teachers believe that the sex differences that they perceive are due to natural causes, i.e. innate factors, or to socialization? If the former, then teachers are more likely to believe that they cannot, or should not, attempt to undermine those differences. Do teachers within a science department, school, or even LEA tend to hold similar views regarding the strengths and weaknesses of boys and girls? An affirmative answer would obviously simplify the introduction of measures to combat sexual inequalities in science departments, and possibly in schools as well.

Finally, some suggestions for future research arose from specific scales. Although the attribution scales failed to confirm the stated hypotheses, it would be unwise immediately to discard the scales at this juncture. The work should be repeated, but instead of referring to boys and girls in general, specific pupils of each teacher should be named and considered individually. Under such conditions, the results may well be different from those obtained in this study. Much more work with the Importance of Subjects scale could profitably be done. For instance, are teachers' views about the importance of different subject areas similar to those of pupils' parents? Such comparisons could lead on to broader issues. Do teachers' views about girls' future adult roles coincide with parents' views and with the views of society in general? Teachers' views concerning a number of educational issues that impinge upon girls' progress in science, e.g. single sex education, could also be compared with parents' views.

10.6 IMPLICATIONS AND RECOMMENDATIONS FOR SCIENCE EDUCATION

The major findings of this study show that science teachers do:

(a) sex type physical science subjects, i.e. they describe them as masculine subjects,

(b) perceive differences between the interests, aptitudes and future roles of girls and boys,

(c) hold higher expectations for boys studying science than for girls,

(d) display sex bias in a marking exercise.

It has been argued (in Chapter 9) that these attitudes, beliefs, expectations and responses are deleterious to girls' advancement in science.

In addition to the broad implications for the science education of girls that can be drawn from the above findings, a number of specific implications also emerged from the study. They mostly highlight the population at which action should be targeted in order to maximise the effectiveness of that action.

1. Work with science teachers.

Science teachers believe that science is of less value to girls than do teachers of other subjects.

2. Work with male science teachers.

Male teachers hold more sex typed ideas about physical science, and hold more sex stereotyped views about people than do female science teachers. In addition, men form the majority of science teachers, and occupy most of the influential positions in the educational system. Work with all-female groups is unlikely to bring about wide ranging and long lasting changes.

3. Work with science teachers in coeducational schools.

Teachers who were working in coeducational schools tended to favour the work of boys over that of girls to a greater extent than did teachers who were teaching in single sex schools. Also, pupils are more inclined to choose sex appropriate subjects in coeducational schools.

4. Work with concrete examples rather than abstract ideas.

Results from the Importance of Subjects scale showed that teachers' views about specific issues were generally less sex differentiated than were their views about abstract issues.

The overall aim of any action directed at science teachers must be to enable them to recognize and modify their sex stereotyped beliefs, attitudes and behaviours. This can best be achieved through pre-service and in-service education. Coverage of sex differences, sex stereotyping and sexism in schools should be a compulsory component of all initial training courses, e.g. BEd, PGCE, for science teachers (and also for teachers of other subjects). At present teacher training institutions tend to introduce these topics in occasional lectures and seminars, if at all (Whyte, 1983a). Practising teachers can be reached via in-service courses, conferences, school based workshops, science teacher centres, and education newspapers, magazines, journals, e.g. The Times Educational Supplement, The School Science Review.

Courses designed to raise teachers' consciousness of the 'Girls and Science' problem and the part that they may play in perpetuating the problem, could focus upon the following objectives.

1. Make teachers aware of sex differentiated patterns of pupil behaviour, subject choice and academic achievement in their own science departments.

2. Make teachers aware of their own sex stereotyped beliefs and attitudes, and their sex differentiated expectations and behaviours.

3. Discuss the possible effects and consequences of such attitudes and behaviours upon girls' prospects in their science lessons.

4. Offer advice and practical suggestions about ways that teachers can improve girls' experience of science.

Although the above is a logical sequence, if the steps were introduced in this order, teachers would probably feel personally affronted and become very antagonistic towards the principles of the course. Nearly all teachers believe that they do not discriminate between their pupils on the basis of sex, although unconsciously and unintentionally they may well do so.

In practice, an order of presentation as outlined below probably stands a greater chance of being sympathetically received by teachers. Thus it also stands a greater chance of effecting change.

1. Alert teachers to the extent of the 'Girls and Science' problem, i.e. their under-representation and under-achievement in science at a number of levels, including employment, and in a number of subjects. Use statistics and facts to impress the reality and gravity of the problem. One teacher has written that "the most difficult stage is accepting the problem and resolving to do something about it" (Hearn, 1979).

2. Discuss possible causes.

Emphasize that 'sex difference' explanations alone are insufficient to explain the differences in educational and vocational outcomes between boys and girls. Stress that very few sex differences are statistically or educationally significant. Ensure compatability between sociological and psychological explanations to enhance their feasibility.

3. Make teachers aware of the part that they may play in dissuading girls from science.

Use games and exercises that involve active participation from the teachers, e.g. the Pupil Preference for Subject Characteristics scale, the Importance of Subjects scale used in this study. The object should be to get teachers to refer and relate to their own circumstances, experiences, actions. Initially, the challenges should occur at the professional, impersonal level, so that teachers do not feel that their personal beliefs are being threatened. But once teachers have questioned their professional beliefs and conduct, it is likely that they will eventually question their personal beliefs as well. In all probability, wide ranging and far reaching changes in teachers' beliefs and responses to girls and science are impossible until they have become committed at the personal level.

4. Offer advice and practical suggestions about ways that teachers can improve girls' experience of science.

Information about various teaching strategies that equalize the treatment of boys and girls is particularly helpful. Teachers may also like to hear about current projects that are attempting to provide equal opportunities for boys and girls. The provision of support or follow-up contact may well prove invaluable.

In conclusion, there is a pressing need to make teachers aware of their sex stereotyped attitudes and behaviours, and to alert teachers to the probable negative outcomes of their beliefs and actions. This cannot be achieved without concerted effort and action. The support and commitment of a large number of bodies will be required, the most obvious being the Association for Science Education, local education authorities, Her Majesty's Inspectorate, national and government agencies, e.g. the Equal Opportunities Commission. A comparatively simple first step would be to press for adequate input about sex differences, sex stereotyping and sexism in schools on all initial training courses for intending science teachers (and intending teachers of other subjects too). These topics should be compulsory, not optional. Furthermore, they should be raised on method courses, where strategies for combating biased teacher behaviour can be discussed in terms of teaching skills. The introduction of sex equity ideas to trainee teachers is likely to be particularly effective as trainee teachers are liable to be more open-minded and impressionable than practising teachers. Also student teachers can carry 'good practice' into large numbers of schools - their teaching practice schools and the subsequent schools to which they are posted.

10.7 CONCLUSION

This study has shown that many science teachers hold attitudes, beliefs and expectations that are likely to adversely affect girls' attitudes towards science and impede their progress in science. It is unlikely that girls' poor level of attainment in science and frequent rejection of science can be explained solely in terms of social, psychological and sex difference factors. The educational process also contributes, and in particular, science teachers themselves. Until now, the attitudes and beliefs guiding science teachers' responses to their male and female pupils have remained largely undocumented. The descriptions of science teachers' views in this thesis will hopefully stimulate and inform future research into teacher effects.

Science teachers who hold sex stereotyped views and sex differentiated expectations help to perpetuate girls' under-achievement in science. Teachers who accept stereotyped assumptions regarding sex differences are likely to hold different expectations for boys and girls and to treat them differently. Because pupils undergo different experiences in the classroom or laboratory based on their sex, they respond differently. Teachers perceive these differences between the behaviour and attainment of boys and girls, and the circle of sexism is complete. However, the circle can easily be broken by changes to just one element. A reduction in teachers' sex stereotyped assumptions, or the equal treatment of boys and girls in science classes could drastically change events. The self-fulfilling prophecy would then operate in a positive manner leading to higher levels of achievement amongst girls.

The role of the individual teacher in helping to improve girls' experience of and gains from their science lessons cannot be overestimated. "When teachers 'change', so does everything in their classrooms" (Frazier & Sadker, 1973, p.xiv). To bring about this change teachers must be made aware of the nature and effects of sexism in the classroom/laboratory, and encouraged to propose ways of combating the problem.

Finally, the findings from the marking exercise should not be ignored. They pose serious questions about the impartiality of science teachers' assessment procedures. Is a marking exercise, which teachers undoubtedly approached very conscientiously, as evidenced by their lengthy comments, totally divorced from usual practices? Further research is urgently required to answer this question. For if only a very slight bias in teachers' marks exists, or if only a proportion of science teachers award sex biased marks, this would still result in many girls receiving prejudiced treatment over the whole country. Biased assessment procedures, together with differential expectations, could be contributing to the personal failure of many girls in science, the minimum of science education for others, and much loss of potential and resources for society.

APPENDICES

Table A1.1/1 The representation of males and females at O level and A level, 1982

(A) O level

	Entries		Passes, grade A-C		
	Boys	Girls	Boys	Girls	Boy:Girl ratio
Physics	134622	49286	81260	30278	2.68 : 1
Chemistry	87074	58651	54860	35247	1.56 : 1
Biology	83610	151061	49556	77171	0.64 : 1

(B) A level

	Entries		Passes, grade A-E		
	Boys	Girls	Boys	Girls	Boy:Girl ratio
Physics	44469	11259	30645	7881	3.89 : 1
Chemistry	30615	16420	22554	12003	1.88 : 1
Biology	17855	25596	12137	17331	0.70 : 1

1.00 : 1.43

Source: DES, Statistics of Education, School Leavers, CSE and GCE 1982

Table A1.1/2 Percentage of entrants obtaining pass grades at O level and A level

			1978	1979	1980	1981	1982
Physics	O level	Boys	58.8	60.1	60.0	59.0	60.4
		Girls	61.0*	61.3*	61.8*	60.7*	61.4*
	A level	Boys	70.7	70.1	70.6*	69.8	68.9
		Girls	71.9*	70.6*	70.5	72.0*	70.0*
Chemistry	O level	Boys	61.3*	62.0*	62.9*	63.0*	63.0*
		Girls	58.7	58.5	59.1	59.7	60.1
	A level	Boys	71.0*	72.2*	71.4	72.9*	73.7*
		Girls	70.4	71.2	72.1*	72.2	73.1
Biology	O level	Boys	58.7*	58.4*	59.9*	59.9*	59.3*
		Girls	52.8	51.5	52.0	52.5	51.1
	A level	Boys	68.4*	67.0*	67.6*	67.7*	68.0*
		Girls	67.5	65.6	67.5	67.5	67.7

* Sex obtaining the higher pass rate

APPENDIX 1.1

THE REPRESENTATION AND ACHIEVEMENT OF GIRLS AND BOYS IN SCIENCE EXAMS

There is no disputing that the representation of girls and boys in external science examinations is unequal. Boys are better represented in the physical science subjects (physics and chemistry), whilst girls are better represented in biology. An inspection of the Department of Education and Science 'Statistics of Education' for a number of years shows very clearly that more boys than girls enter and pass physics and chemistry at O level and A level in England and Wales. More girls than boys enter and pass biology in these examinations. Table A1.1/1 uses the most recent figures available to illustrate the ratio of boys to girls currently entering and passing science subjects at the two examination levels.

A detailed comparison of the percentage of entrants obtaining pass grades reveals that boys have consistently achieved a higher pass rate in chemistry and biology at both O and A level over the past five years. However, girls have consistently achieved a higher pass rate in physics (see Table A1.1/2). This observation seems not to accord with the commonly held view that girls under-achieve in the physical science subjects.

The higher pass rate of girls in physics can be attributed to the high degree of selection that preceded their success. Since comparatively few girls attempt physics exams, they are likely to be atypical. Having chosen physics against a sex barrier (the ratio of boys to girls in physics is more extreme than for any other science subject), they are likely to be more motivated, more dedicated to their studies and more intelligent than the larger group of girls who choose the other physical science subject - chemistry. Thus girls' better performance in physics than in chemistry is not altogether surprising.

Although the few girls who actually attempt physics achieve a higher overall percentage pass rate than the boys, yet the boys achieve a higher

Table A1.1/3 Percentage of successful O level entrants (grades A-C) who obtained grade A in 1982

	Boys	Girls
Physics	20.2	17.3
Chemistry	20.7	16.8
Biology	21.0	16.2

Table A1.1/4 Percentage of successful A level entrants (grades A-E) who obtained grades A and B in 1982

	Boys	Girls
Physics	35.7	32.4
Chemistry	39.0	35.0
Biology	34.4	32.4

percentage of good passes. Of the pupils who pass O level physics, a higher proportion of boys than girls obtain a grade A. Similarly at A level, of the pupils who pass physics, a higher proportion of boys than girls obtain a grade A or B, which are the grades most likely to secure university entrance. This trend of boys being more successful at achieving the higher grades can be detected over a number of years at O and A level physics results, and also in other science subjects as well (see Tables A1.1/3 and A1.1/4).

The assertion that girls under-achieve in the physical science subjects is vindicable. In physics and chemistry at both O level and A level, boys are more successful at achieving the better pass grades than are girls. Furthermore, boys achieve a higher overall pass rate in chemistry than do girls. Taking boys' and girls' entries for physics and chemistry examinations into account, then not only do more boys attempt physical science exams, but they also achieve better results. The fact that a higher proportion of girls than boys pass physics exams should not detract from the poorer performance of the most able girls. As more girls enter physics exams, it is likely that their overall pass rate will decline. There is still every justification for our concern over the numbers of girls attempting the physical science subjects, and their level of achievement in physical science examinations.

APPENDIX 4.1

VARIANCE CONTROL

The control of variance is a major objective of research design. Variance control is sought by maximising the variance of the variables under investigation, by minimising or isolating the variance of variables which are extraneous to the purposes of the research, and by minimising error or random variance.

The variance of the dependent variable influenced by the independent variable or variables of an investigation (the experimental variance) should be maximised, so that it is discernible from the total variance of the dependent variable. The use of independent variables which vary substantially helps to maximise experimental variance. Davis (1971) has formulated this principle into a rule, "Make sure your variables vary" (p.23). As a result of paying due regard to this rule, some variables in the Personal Details section of the BIAS questionnaire were discarded at the pilot stage.

Systematic variance due to extraneous, independent variables should be minimised or separated from the variance of the independent variables under investigation. There are a number of ways of controlling extraneous variables. One method is to eliminate the effect of an independent variable by choosing subjects so that they are as homogeneous as possible on the variable. This method was not purposely adopted since it severely restricts the generalizability of the results. However, in all probability secondary school science teachers, the population for the main body of this research, are comparatively homogeneous on a number of variables. A second and preferred way is by randomization. Ideally subjects should be selected at random and the different measurement instruments should be allocated at random. In this research randomization was aimed for but was not necessarily achieved. A third method of controlling an extraneous variable is by including it in the research design as an attribute variable. It is then possible to extract the variance due to the variable from the total variance of the dependent variable. Besides providing information about the effect of the variable on the dependent variable, this method also allows possible interaction with other independent variables to be investigated. These considerations governed the inclusion of a number of the variables appearing in the School Details questionnaire and the Personal Details section of the BIAS questionnaire in the present study. A fourth way of controlling extraneous variance is to match subjects. This method was not attempted in this study. Another form of control arises from the use of statistical methods. Statistical control can isolate and quantify variance.

Error variance is the variability of measures due to random instability. Although random errors tend to cancel each other our, error variance is fundamentally unpredictable. Error variance should be minimized to allow systematic variances under investigation to be apparent and to reach statistical significance. Sources of error variance include certain factors associated with individual differences between subjects and also measurement errors, for example, variation of responses over time, guessing, transient lapses of concentration. The two commonest approaches to minimizing error variance concentrate upon improving measurements by reducing errors and increasing reliability. In this study attempts were made to reduce errors of measurement by issuing clear, detailed instructions to respondents. In addition, superfluous measurement scales, which although theoretically interesting were not closely related to the research problem under investigation, were discarded during the early stages of the research. Error variance was also reduced by paying due attention to the reliability of the measurement scales used (section 6.3). Increasing the reliability of a measure, has the effect of reducing the error variance in relation to the total variance.

APPENDIX 4.2

ADVANTAGES AND DISADVANTAGES OF THE RESEARCH METHODS USED

4.2.1.A Exploratory interviews

When starting work in a previously under-researched area, an important first task is to specify and clarify the issues and problems. One method is to ask subjects directly about their views on the issues. Assuming that the subjects have or can formulate opinions on the topics and that they are prepared to voice those opinions, rather than ones that they think the researcher would like to hear, then direct questioning can produce much useful and relevant information. Exploratory interviews can help to identify significant variables and relationships among those variables, suggest hypotheses for later investigation, and guide the development of measuring instruments.

Interviews have certain advantages that make them suitable for exploratory work. The interviewer can clarify questions to ensure that the respondent understands them, and can also probe into the context and reasons for answers to particular questions. Rapport can be built up and maintained in order to ensure that the respondent is kept interested and responsive to the end of the interview. This tactic should also encourage the respondent to give spontaneous, and therefore probably more truthful, replies. The spontaneity and richness of information collected by interviews can provide good illustrations of attitudinal and perceptual stances (see Chapter 5).

Disadvantages of interviews are that they are costly, since travelling is involved, and they are very time consuming. Getting information from one subject may take up to an hour. In this study these problems of time and cost were offset by interviewing a comparatively small sample of science teachers from only a few schools. Other disadvantages revolve around bias effects. The interviewer's characteristics and behaviour can influence the replies expressed by a respondent (Moser & Kalton, 1971). Also the interviewer's expectations and selective understanding of the answers may bias the record made of the interview (Oppenheim, 1966). No special attempt was made to suppress the biasing influence of the interviewer in the exploratory interviews conducted, beyond her trying to adopt a reasonably impartial role. To counter biased recording, the exploratory interviews were recorded on tape and then transcribed before being analysed.

4.2.2.A Field experiment

A field experiment is an experiment which takes place in a natural setting. The researcher is able to manipulate one or more independent variables under as carefully controlled conditions as the situation will allow. Field experiments are particularly well suited to the testing of theory and hypotheses in the area being researched.

The results obtained from field experiments are often weakened by contamination from uncontrolled environmental variables. In addition, the variables under investigation can be diluted if subjects consult each other during the course of the experiment. In the marking exercise, teachers were specifically told not to discuss the exercise until they had finished marking it. The effect of the manipulated variables on the dependent variable can also be maximised by measuring the dependent variable as precisely as possible. Unfortunately dependent variable measures of the type studied in this research are often rather insensitive (Kerlinger, 1973). In the marking exercise 5-point scales were used to measure the dependent variables.

4.2.3.A Questionnaires

Questionnaires are often used when a researcher wants answers to a variety of questions, for questionnaires are well adapted to obtaining personal and social facts, beliefs and attitudes. However, there is always an underlying assumption that the questions have the same meaning to all the respondents, resulting in comparability of responses.

The chief advantage of the questionnaire method is its ability to collect a great deal of information from a large sample in a relatively short period of time. Questionnaire surveys are also efficient because the cost of collecting data is low. If they are posted, low costs can still be combined with wide geographical coverage. Questionnaires generally permit anonymity. This should increase the chances of receiving responses that genuinely reflect the respondents' views and beliefs. Another advantage is that the questions are standardized, so ensuring a considerable degree of uniformity across measurement situations. Each subject responds to exactly the same questions. Standardized questions produce data that is easily processed and that is particularly amenable to statistical analysis, thus aiding interpretation of the results.

The chief disadvantage of postal questionnaires is the fact that they usually produce very poor response rates. This is a problem, not because of reduced sample size, but because of the possibility of a biased sample resulting, from which generalizations could not be made. There are techniques available for increasing postal returns, e.g. reminder letters (Parten, 1950), and they were employed in this study. Another technique to encourage high returns is to keep the questionnaire brief. However, this constraint can be seen as a disadvantage. In this study it was thought prudent to shorten the questionnaires between the pilot forms and the final forms. Besides being relatively brief, postal questionnaires must also be simple, to ensure that they are completed correctly. This means that questionnaires are better suited to extensive research rather than intensive research (Kerlinger, 1973). However, there is no inherent reason why questionnaires cannot be used to investigate sociological and psychological variables in depth, and this was attempted in the present study. Another disadvantage associated with the structure of questionnaires is that they are inflexible. The survey method using postal questionnaires can be weak on control. Since respondents fill questionnaires on their own, they may misunderstand questions, omit questions, or consult other people. Moreover, it is generally impossible to check the reliability of the responses given. In this research it was hoped that careful wording of questions and detailed instructions would minimise errors from lack of control. Finally, it must be remembered that the return of questionnaires from the last few respondents can be very slow, and so the survey must be spread over a period of months if a reasonable response rate is to be achieved.

APPENDIX 4.3

FACTORS INFLUENCING CHOICE OF PRIMARY RESEARCH METHOD

The decision to collect most of the data in this investigation by postal questionnaire was guided by a number of factors, which can conveniently be grouped under three major headings: resources, constraints and need for large samples.

4.3.1.A Resources

(a) Existing contacts with a very limited number of secondary schools scattered throughout England.

(b) Contact with a number of people possessing considerable expertise in the construction of questionnaires and the planning of surveys.

(c) The opportunity of having a reprographic department reproduce material professionally and speedily.

(d) Access to computing facilities and the availability of expert guidance and tuition in their use.

4.3.2.A Constraints

(a) Limited period of time in which to complete the research.

(b) Limited manpower - basically one person.

(c) Difficulties associated with obtaining funds for extensive travelling.

(d) Teachers are very busy people, so it is difficult to arrange to see them.

(e) Schools are prepared to agree to a single visit from a researcher, but they are much less inclined to agree to a long term project.

4.3.3.A Need for large samples

A number of the research requirements listed below could best be satisfied by drawing large samples from over the whole of England.

(a) The desire to investigate not only the existence of sex bias and sex stereotyping amongst science teachers, but also to record how widespread these views and practices are.

(b) The need to sample sufficient numbers of rare individuals, e.g. female physics teachers.

(c) The desire to explore possible relationships between the central topics and a number of independent variables, relating not only to the subjects, but also to the schools in which they teach.

(d) It is intended that the main findings of the study will be widely disseminated to science teachers. Having received extensive training in the 'hard' sciences, some science teachers are sceptical of the research methods used in the social sciences. Nevertheless, it was felt that science teacher would be more likely to accept the representativeness of a large sample and to accept generalizations from it.

(e) Since computing facilities were available, there was the opportunity of performing detailed statistical analyses. This option was welcomed as it was felt that science teachers would view favourably findings that were supported by statistical arguments.

4.3.4.A Additional considerations

(a) I felt that the use of ethnographic methods would not be feasible, as I doubted my ability to view the familiar as strange and problematical. Having been a science teacher myself, I tend to identify very closely with the teacher's role and to accept the status quo.

(b) Extensive interviewing was rejected because it was feared that experimenter bias could unduly influence the analysis of the data.

(c) Reactive effects could influence the results obtained from interviews. Subjects might sense my attitudes towards the topics under discussion and modify their own attitudes accordingly, or they might temper their statements simply because I am a female.

Consideration of the above points suggested that questionnaires would be the most appropriate method of gathering data. Furthermore, it was decided to distribute these questionnaires by post to facilitate contacting the number of schools whose assistance would be required during the pilot stages, and also to facilitate contacting a large sample of teachers from schools of differing characteristics during the data collection stage.

APPENDIX 4.4

THE BIAS QUESTIONNAIRE

BASES OF INDIVIDUAL ASSESSMENT

Teachers systematically assess all aspects of their pupils' work and progress. But different teachers use different criteria and so may arrive at different judgements and recommendations. By studying the importance attached to a variety of work and pupil characteristics, it is hoped that the under-achievement of certain groups of pupils in science may be better understood and perhaps remedied.

All replies will be treated as strictly confidential and no teachers or schools will be referred to individually in any ensuing report.

The first section of the questionnaire requests information regarding your educational and teaching experiences. This data is required solely for classification purposes. The second section entails marking samples of pupils' work.

The design of the investigation requires that you do not confer with your colleagues until you have completed the questionnaire.

Please answer all the questions as incomplete returns will seriously limit subsequent analysis.

Any comments that you would like to make regarding the aims and/or design of this investigation will be most welcome.

Finally, thank you for giving me your help and time.

Margaret Spear
IET, The Open University
Autumn 1981

SECTION A

This section refers largely to your educational and teaching experiences. Most of the questions require you to place a tick in the relevant box. Please ensure that you answer all the questions.

1. Your name ____________________

2. Name of present school ____________________

3. Sex: Male ☐ Female ☐

4. Age: Under 30 ☐
   30 - 39 ☐
   40 - 49 ☐
   50 and over ☐

5. Background:
   (a) What is the best classification of the household in which you spent your childhood?
   Working class ☐
   Middle class ☐
   Upper class ☐
   (b) Was your mother a full-time housewife for the majority of the time you were a child? Yes ☐ No ☐

6. Type(s) of school attended during your secondary education:
   (Answer parts (a) and (b). Tick as many boxed in both parts as appropriate)
   (a) Secondary modern ☐
   Grammar ☐
   Comprehensive ☐
   Direct grant ☐
   Independent ☐
   Other (specify) ☐ ____________________
   (b) Co-educational ☐
   Single sex: boys' ☐
   girls' ☐

7. Qualifications:
   Bachelor of Education ☐
   Untrained graduate or equivalent ☐
   Graduate with teaching qualification or equivalent ☐
   Certificated teacher ☐
   Other (specify) ____________________ ☐

8. Teaching experience:
   Less than 2 years ☐
   2 - 5 years ☐
   5 - 10 years ☐
   10 - 20 years ☐
   Over 20 years ☐

9. Current principal teaching subject ____________

10. Age of youngest pupils you are teaching any subject to at present ______

11. Which of the following subjects have you taught (for at least a year)?

- Physics ☐
- Chemistry ☐
- Biology ☐
- Integrated Science ☐

12. Have you ever taught in a school in which a balanced science curriculum (one comprising the basic elements of biology, chemistry and physics) was compulsory for all pupils up to the age of 16? Yes ☐ No ☐

13. Which of the following types of syllabus have you taught (for at least a year)?

CSE
- Mode 1 ☐
- Mode 2 ☐
- Mode 3 ☐

GCE 'O' level
- Traditional ☐
- Nuffield ☐
- SCISP ☐

GCE 'A' level
- Traditional ☐
- Nuffield ☐

14. In which of the following types of school have you taught? (Exclude teaching practice) (Tick as many boxes in both parts as appropriate)

(a) Secondary modern ☐
- Grammar ☐
- Comprehensive ☐
- Direct grant ☐
- Independent ☐
- Primary ☐
- Other (specify) ☐ ____________________

(b) Co-educational ☐
- Single sex: boys' ☐
- girls' ☐

SECTION B

On the following pages you will see six samples of work produced by three pupils of average ability, from the same class of a comprehensive school. The work was produced shortly after the children entered secondary school at the age of twelve. The children were following a Combined Science course taught in mixed ability groups. Although you may not be teaching Combined Science yourself, your qualifications and experience are appropriate for evaluating these samples.

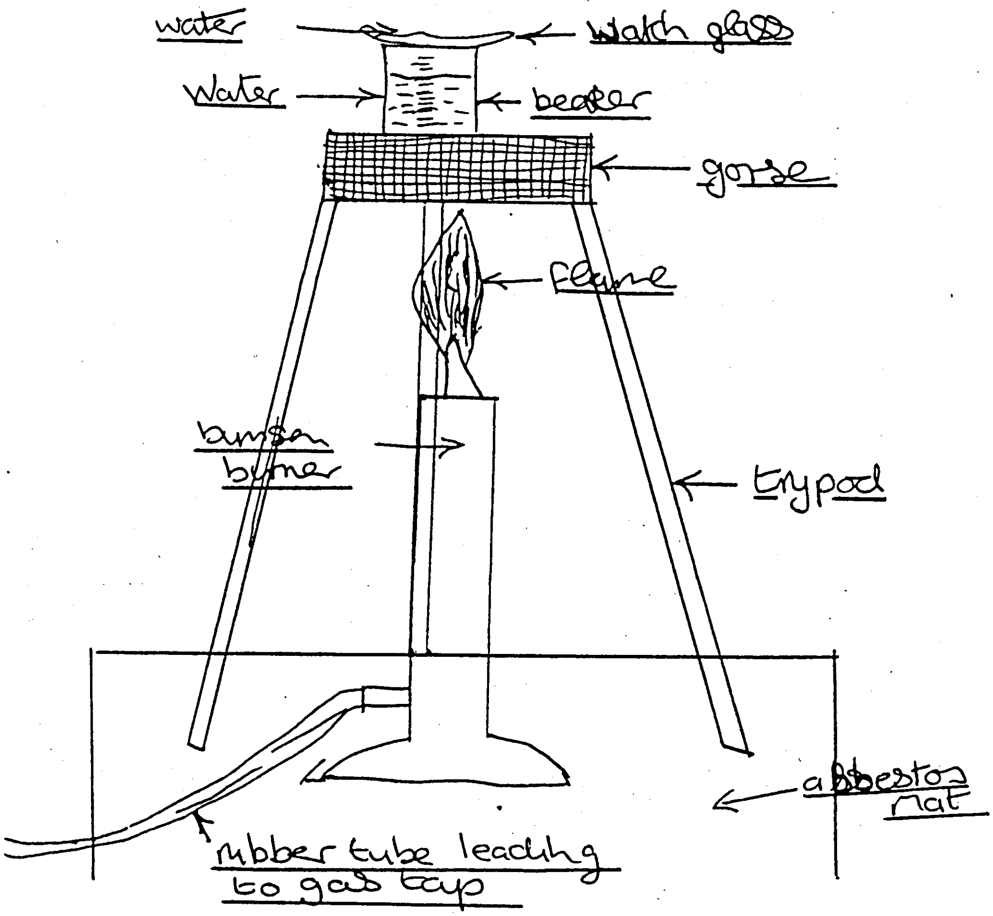
The first sample from each child is their write up of a practical experiment on distillation. The class were first introduced to the ideas of physical change and change of state. Then after a brief discussion about the process of distillation, they were presented with the practical problem of producing a sample of distilled water from tap water using only simple apparatus. After selecting their apparatus and carrying out the distillation, the pupils had to write up their experiment.

The second sample from each child is an essay which was set for homework. The title was 'What I think about science and scientists'.

Claire

Distillation of Water

The water was boiled until it had evaporated and turned to steam. The watch glass prevented the steam escaping. The flame was then removed and the steam was left to cool. When the ~~water~~ steam had cooled the water was distilled water.



Claire

What I think of Science and Scientist

① a) A scentist is a person who mixes dangerous chemicals to other dangerous chemicals sometimes he discoveres things wich are good and help people who are ill and who have uncurable diseases until he finds something which will cure this decease but sometimes he discovers bad things which could kill somebody.

b) I have told what he does to suntane extent as well as making good and bad things he also carries on research about things already found and helps to make them better.

c) A scentist works in a laboratory where they have all kinds of instruments, test tubes, lights, ~~___~~ chemicals, eye peaces, microscopes and so on.

d) He investigates what chemicals do to other chemicals and good and bad things. You would find laboratries in colleges special places made for laboratries you could find them in factories as well but that would be for the things the factory makes, eg. clothes, Metals, cars, machinery, etc.

② Science is about all of ~~of~~ things like plants and animals it diseases about new ways of communicating per animals. Its about new inventions like the cillicon chip, atomic bombs and things like that.

DISTILLATION OF WATER by CLAIRE

1. From your experience, is the standard of this girl's work

   above average ☐
   average ☐
   below average ☐

2.(a) What mark out of 10 does it merit? ☐

  (b) What mark out of 10 would you actually give it? ☐
  (The answer to parts (a) and (b) may not necessarily be the same. They could differ for a variety of reasons, e.g. you may wish to motivate the child to improve her work.)

3. How would you rate this girl's work on the following factors?
   Place a tick in the appropriate box to show your assessment of each factor.

	Very good				Very poor
List of factors	1.	2.	3.	4.	5.
1. Neatness					
2. Effort involved					
3. Grammar & spelling					
4. Scientific accuracy					
5. Understanding of principles . . .					
6. Clarity of explanation					
7. Standard of diagram					

4. How would you judge this girl's aptitude for science?
   (Tick appropriate space along the scale)
   Considerable aptitude ___ : ___ : ___ : ___ : ___ Little aptitude

WHAT I THINK OF SCIENCE AND SCIENTIST by CLAIRE

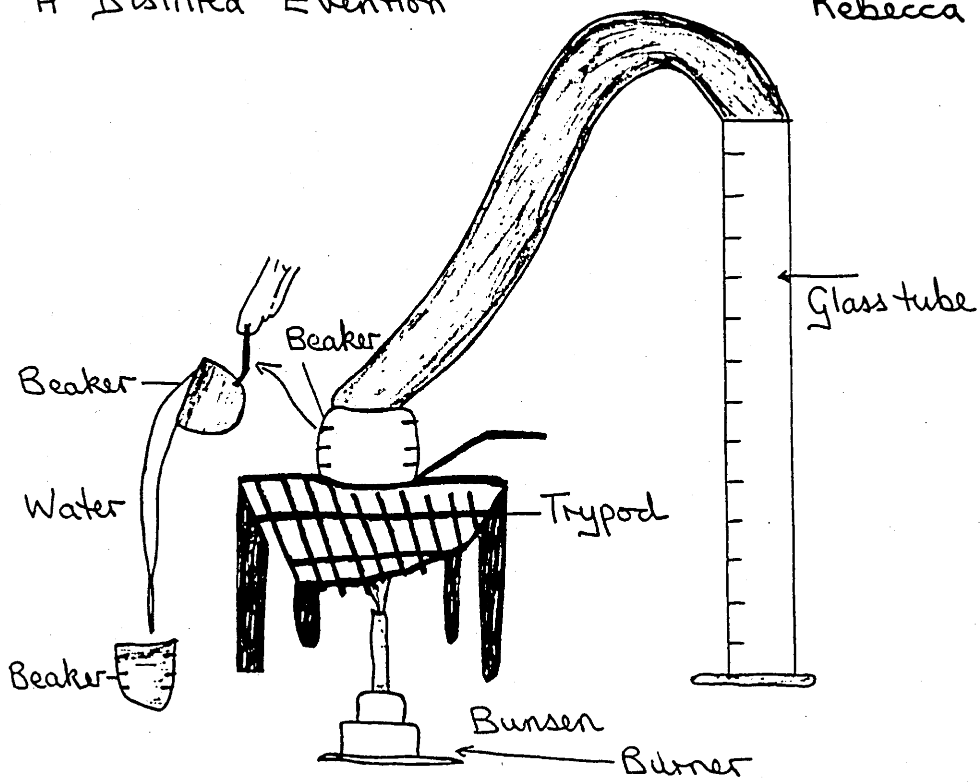
1. How would you judge this girl's attitude towards science?
   Favourable attitude ___ : ___ : ___ : ___ : ___ Unfavourable attitude

2. How would you judge this girl's interest in science?
   Very interested ___ : ___ : ___ : ___ : ___ Not interested

3. How would you rate this girl's eventual suitability for 'O' level physical science courses?
   Highly suitable ___ : ___ : ___ : ___ : ___ Highly unsuitable

4. How would you rate this girl's eventual suitability for CSE physical science courses?
   Highly suitable ___ : ___ : ___ : ___ : ___ Highly unsuitable

A Disitted Evention Rebecca



Water is contained in this big glass test tube which is tipped up and a certain amount of water flows through a plastic rubber tube and into a beaker. This is placed on a trypod and underneath the trypod is a bunsen Burner. This is lit and the heat comes through the tripod and heats up the water in the beaker and when the water is hot enough it is tipped up and poured into another beaker and is ready to use as distilled water.

What I think about Science and Scientists   Rebecca

A scientist is someone who finds out more about the world around us. A scientist can work in all sorts of places, ~~he ca~~ a scientist can work outside discovering more about the world and its living things. A scientist can work inside doing different experiments on different things. A scientist usually works hard to find out about the world though it would be quite impossible for a scientist to know everything about the world.

Science is nearly everything around us. Science is finding out about things. Science is discovering things and trying out new things eg sending man to the moon. Science ~~of~~ is finding out more about space and the stars. I think that a dog was sent up to space first, I do not think that they should try science out on animals as the animals ~~sh~~ live ~~less~~ years than what we do so why should they be tested on. Science affects our lives ~~a~~ a lot.

A DISLITTED EVENTION by REBECCA

1. From your experience, is the standard of this girl's work
   - above average ☐
   - average ☐
   - below average ☐

2.(a) What mark out of 10 does it merit? ☐

(b) What mark out of 10 would you actually give it? ☐
(The answer to parts (a) and (b) may not necessarily be the same. They could differ for a variety of reasons, e.g. you may wish to motivate the child to improve her work.)

3. How would you rate this girl's work on the following factors?
Place a tick in the appropriate box to show your assessment of each factor.

	Very good				Very poor
List of factors	1.	2.	3.	4.	5.
1. Neatness					
2. Effort involved					
3. Grammar & spelling					
4. Scientific accuracy					
5. Understanding of principles . . .					
6. Clarity of explanation					
7. Standard of diagram					

4. How would you judge this girl's aptitude for science?
(Tick appropriate space along the scale)

Considerable aptitude ___ : ___ : ___ : ___ : ___ Little aptitude

WHAT I THINK ABOUT SCIENCE AND SCIENTISTS by REBECCA

1. How would you judge this girl's attitude towards science?

   Favourable attitude ___ : ___ : ___ : ___ : ___ Unfavourable attitude

2. How would you judge this girl's interest in science?

   Very interested ___ : ___ : ___ : ___ : ___ Not interested

3. How would you rate this girl's eventual suitability for 'O' level physical science courses?

   Highly suitable ___ : ___ : ___ : ___ : ___ Highly unsuitable

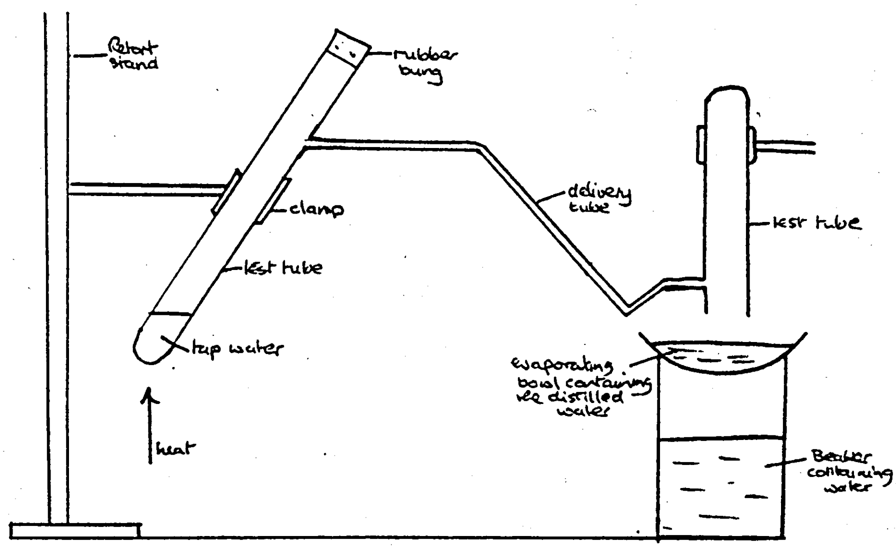
4. How would you rate this girl's eventual suitability for CSE physical science courses?

   Highly suitable ___ : ___ : ___ : ___ : ___ Highly unsuitable

Matthew

My apparatus for separating pure water from tap water

The first apparatus we set up was what seemed to us good but when a ~~si~~ simular one was tried out the one problem was that after it had been heated up the rubber corke would shoot off. So after a little ~~white~~ thinking we found that it needed some air outlet. But with our second idea when the steam got to one point it all got away. So ~~then~~ we had to think of an apparatus where the stopper would not blow off and no steam could escape.



Matthew

What I think about science and scientists

A scientist is a man who discovers cures for illnesse's and does experiments to find out different chemicals to improve the human life. The thing I don't like about scientists is that they use their experiments on little animals like guinea pigs and rats etc. Mind you what else is there to, experiment on. If the scientists didn't use their experiments on animals they'd most probably use volunteered humans. A scientist would most probably be found working at a University. But a top scientist would most probably never be found by quite a lot of people, only the people who the top scientist works for. I think the best thing the scientist have ever made is medicines because they've saved millions and millions of lives. Science is a way of finding out about life around us and ~~f~~ experimenting on cures and things. Science is a very important thing in our lives.

MY APPARATUS FOR SEPARATING PURE WATER FROM TAP WATER by MATTHEW

1. From your experience is the standard of this boy's work
   - above average ☐
   - average ☐
   - below average ☐

2.(a) What mark out of 10 does it merit? ☐

(b) What mark out of 10 would you actually give it? ☐

(The answer to parts (a) and (b) may not necessarily be the same. They could differ for a variety of reasons, e.g. you may wish to motivate the child to improve his work.)

3. How would you rate this boy's work on the following factors?
Place a tick in the appropriate box to show your assessment of each factor.

List of factors	Very good 1.	2.	3.	4.	Very poor 5.
1. Neatness					
2. Effort involved					
3. Grammar & spelling					
4. Scientific accuracy					
5. Understanding of principles . . .					
6. Clarity of explanation					
7. Standard of diagram					

4. How would you judge this boy's aptitude for science?
(Tick appropriate space along the scale)

Considerable aptitude ___ : ___ : ___ : ___ : ___ Little aptitude

WHAT I THINK ABOUT SCIENCE AND SCIENTISTS by MATTHEW

1. How would you judge this boy's attitude towards science?

Favourable attitude ___ : ___ : ___ : ___ : ___ Unfavourable attitude

2. How would you judge this boy's interest in science?

Very interested ___ : ___ : ___ : ___ : ___ Not interested

3. How would you rate this boy's eventual suitability for 'O' level physical science courses?

Highly suitable ___ : ___ : ___ : ___ : ___ Highly unsuitable

4. How would you rate this boy's eventual suitability for CSE physical science courses?

Highly suitable ___ : ___ : ___ : ___ : ___ Highly unsuitable

APPENDIX 4.5

THE STOSS QUESTIONNAIRE

SCIENCE TEACHERS ON SCIENCE SUBJECTS

The main purpose of this survey is to record the views of science teachers on (a) the subjects they teach
and (b) their pupils.

Although much is known about pupils' attitudes towards the science subjects, little is known about how teachers view the subjects that they teach. I hope to record these views and then relate them to whether teachers regard science to be equally suitable for all types of pupils. The results will be used to assess teachers' attitudes towards compulsory science courses for all pupils up to 16.

All replies will be treated as strictly confidential and no teachers or schools will be referred to individually in any ensuing report.

The design of the survey requires that you do not confer with your colleagues until you have completed the questionnaire.

Please answer all the questions as incomplete returns will seriously limit subsequent analysis.

I hope that completing this questionnaire will be an interesting exercise for you. Any comments that you would like to make will be most welcome.

Margaret Spear
IET, The Open University

Your name for reference purposes only ________________

Over the past few years, what has been your principal teaching subject?

- Physics ☐
- Chemistry ☐
- Biology ☐
- Integrated Science ☐
- Other (specify) ☐ ____________

(Tick ONE box only)

In this questionnaire the phrase 'YOUR SUBJECT' refers to the subject you have ticked above.

SECTION A

This section is designed to explore science teachers' feelings about science subjects as they are taught up to CSE/O level standard. Each subject being investigated is followed by a series of adjective pairs. You are asked to place a tick on one of the seven lines between each pair of adjectives to show how you feel personally about the subject.
For example, if you think that Physics, the first subject listed below, is extremely hard you would place a tick on the extreme left-hand line.

Hard ✓ : ___ : ___ : ___ : ___ : ___ : ___ Soft
extremely quite slightly neutral slightly quite extremely

If you believe that Physics is extremely soft, put a tick on the extreme right-hand line.

Hard ___ : ___ : ___ : ___ : ___ : ___ : ✓ Soft
extremely quite slightly neutral slightly quite extremely

If your ideas about Physics are not so extreme, place a tick in the appropriate intermediate position.
It is your immediate response to each item that is required, even if some of the scales may seem inappropriate to describe science subjects. Therefore trust your first impressions and work quickly.
Do not attempt to compare your responses for the three subjects. Each tick should represent a separate and independent judgement.
PLEASE BE SURE YOU TICK EVERY SCALE FOR ALL THREE SUBJECTS.

PHYSICS

Hard ___ : ___ : ___ : ___ : ___ : ___ : ___ Soft
Tender ___ : ___ : ___ : ___ : ___ : ___ : ___ Tough
Cold ___ : ___ : ___ : ___ : ___ : ___ : ___ Warm
Intimate ___ : ___ : ___ : ___ : ___ : ___ : ___ Remote

CHEMISTRY

Hard ___ : ___ : ___ : ___ : ___ : ___ : ___ Soft
Tender ___ : ___ : ___ : ___ : ___ : ___ : ___ Tough
Cold ___ : ___ : ___ : ___ : ___ : ___ : ___ Warm
Intimate ___ : ___ : ___ : ___ : ___ : ___ : ___ Remote

BIOLOGY

Hard ___ : ___ : ___ : ___ : ___ : ___ : ___ Soft
Tender ___ : ___ : ___ : ___ : ___ : ___ : ___ Tough
Cold ___ : ___ : ___ : ___ : ___ : ___ : ___ Warm
Intimate ___ : ___ : ___ : ___ : ___ : ___ : ___ Remote

SECTION B

Do you believe that the following two science subjects, as they are taught in secondary schools up to CSE/O level standard, possess any of the following characteristics? Please indicate how logical, objective, etc. you think both of the subjects are, by placing a tick in the appropriate column for each item.

PHYSICS	Very	Fairly	Not very	Not at all
1. Logical				
2. Objective				
3. Relevant for careers				
4. Relevant for family life				
5. Mathematical				
6. Wordy				
7. Concerned with people				
8. Concerned with objects				
9. Concerned with social issues				
10. Unfamiliar				
11. Technical				
12. Mechanical				
13. Masculine				
14. Abstract				
15. Impersonal				

BIOLOGY	Very	Fairly	Not very	Not at all
1. Logical				
2. Objective				
3. Relevant for careers				
4. Relevant for family life				
5. Mathematical				
6. Wordy				
7. Concerned with people				
8. Concerned with objects				
9. Concerned with social issues				
10. Unfamiliar				
11. Technical				
12. Mechanical				
13. Masculine				
14. Abstract				
15. Impersonal				

<u>SECTION C</u>

What characteristics do you believe make a subject attractive to 14 year old pupils? Using the 7-point scales below, place ticks to show the extent to which <u>both</u> girls and boys prefer the different characteristics.

A typical 14 year old <u>girl</u> prefers subjects which are:

Theoretical	___	___	___	___	___	___	___	Practical
Numerical	___	___	___	___	___	___	___	Verbal
Science	___	___	___	___	___	___	___	Arts
Logical	___	___	___	___	___	___	___	Intuitive
Feminine	___	___	___	___	___	___	___	Masculine
Factual	___	___	___	___	___	___	___	Opinionative
Creative	___	___	___	___	___	___	___	Routine
Simple	___	___	___	___	___	___	___	Complex
Important	___	___	___	___	___	___	___	Unimportant

A typical 14 year old <u>boy</u> prefers subjects which are:

Theoretical	___	___	___	___	___	___	___	Practical
Numerical	___	___	___	___	___	___	___	Verbal
Science	___	___	___	___	___	___	___	Arts
Logical	___	___	___	___	___	___	___	Intuitive
Feminine	___	___	___	___	___	___	___	Masculine
Factual	___	___	___	___	___	___	___	Opinionative
Creative	___	___	___	___	___	___	___	Routine
Simple	___	___	___	___	___	___	___	Complex
Important	___	___	___	___	___	___	___	Unimportant

SECTION D

Pupils of 13 - 14 years commonly have to choose a number of subject options. This section considers some of the reasons which can influence pupils to either accept or reject a particular subject.

1. What factors do you think influence those girls who choose YOUR SUBJECT? (Remember YOUR SUBJECT refers to the subject you ticked on the first page) Indicate how frequently you believe each of the following reasons apply.

Reasons	Never applies	Rarely applies	Sometimes applies	Often applies	Always applies
Find subject easy					
Like subject content					
Like teaching style					
Expect to like teaching group .					
Like teacher					
Relevancy for future career . .					
Relevancy for future family life					
Parental influence					
Tradition					
Girls' subject					

2. What factors do you think influence those girls who drop YOUR SUBJECT? Indicate how frequently you believe each of the following reasons apply.

Reasons	Never applies	Rarely applies	Sometimes applies	Often applies	Always applies
Find subject difficult					
Dislike subject content					
Dislike teaching style					
Expect to dislike teaching group					
Dislike teacher					
Irrelevancy for future career .					
Irrelevancy for future family life					
Parental influence					
Tradition					
Boys' subject					

SECTION E

This section considers possible reasons for pupils' academic success and failure in science subjects up to CSE/O level standard.

1. What do you think are the most important factors contributing to girls' success in YOUR SUBJECT area?
   Please indicate how important you believe each of the following reasons to be.

Reasons	Very important	Fairly important	Not very important	Not at all important
Ability				
Out-of-class experience				
Good attitude				
Motivation				
Effort				
Interest in subject				
Conscientiousness				
Attentiveness				
Good teaching				
Subject simplicity				
Assistance from peers				
Family support				
Emotional stability				

2. What do you think are the most important factors contributing to girls' failure in YOUR SUBJECT area?
   Please indicate how important you believe each of the following reasons to be.

Reasons	Very important	Fairly important	Not very important	Not at all important
Lack of ability				
Little relevant out-of-class experience . . .				
Poor attitude				
Lack of motivation				
Lack of effort				
Lack of interest in subject . .				
Carelessness				
Lack of attention				
Poor teaching				
Subject difficulty				
Distraction by peers				
Lack of family support				
Lack of emotional stability . .				

SECTION F

How do you picture scientists from the different disciplines?
Tick the appropriate line on each scale.

a PHYSICIST is

Male								
probable	___	___	___	___	___	___	___	improbable
Good at maths								
probable	___	___	___	___	___	___	___	improbable
Logical								
probable	___	___	___	___	___	___	___	improbable
Objective								
probable	___	___	___	___	___	___	___	improbable
Competitive								
probable	___	___	___	___	___	___	___	improbable
Unsociable								
probable	___	___	___	___	___	___	___	improbable
Unemotional								
probable	___	___	___	___	___	___	___	improbable
Humanitarian								
probable	___	___	___	___	___	___	___	improbable

a BIOLOGIST is

Male								
probable	___	___	___	___	___	___	___	improbable
Good at maths								
probable	___	___	___	___	___	___	___	improbable
Logical								
probable	___	___	___	___	___	___	___	improbable
Objective								
probable	___	___	___	___	___	___	___	improbable
Competitive								
probable	___	___	___	___	___	___	___	improbable
Unsociable								
probable	___	___	___	___	___	___	___	improbable
Unemotional								
probable	___	___	___	___	___	___	___	improbable
Humanitarian								
probable	___	___	___	___	___	___	___	improbable

SECTION G

Listed below are some statements describing attitudes towards the role of women in society. Please indicate your response to each statement by placing a circle around the option which most accurately reflects your opinion. The response options are:

AA - strongly agree
a - agree
? - no opinion
d - disagree
DD - strongly disagree

1. As head of the household, the father should have final authority over his children.	AA	a	?	d	DD
2. A woman who refuses to give up her job to move with her husband would be to blame if the marriage broke up.	AA	a	?	d	DD
3. A woman who refuses to bear children has failed in her duty to her husband.	AA	a	?	d	DD
4. A woman should be expected to change her name when she marries.	AA	a	?	d	DD
5. It is all right for women to work but men will always be the basic breadwinners.	AA	a	?	d	DD

Thank you for your cooperation and assistance.

Margaret Spear

APPENDIX 4.6

THE COSS QUESTIONNAIRE (For secondary school teachers)

CHARACTERISTICS OF SCHOOL SUBJECTS

The purpose of this survey is to investigate your views about several school subjects as they are taught in secondary schools up to CSE/O level standard. It forms part of a larger and more detailed study linking teachers' attitudes towards different subject areas with the encouragement and advice that they offer to different groups of pupils. Your replies to this questionnaire will greatly help the interpretation of future research.

PLEASE SUPPLY

(For statistical purposes only)

Sex: Female ☐ Male ☐

Age: Under 30 ☐ 40 - 49 ☐

30 - 39 ☐ 50 and over ☐

Your principal teaching subject ____________________

Subsidiary teaching subject(s) ____________________

DIRECTIONS

On the following pages you will find a selection of school subjects and beneath each a set of descriptive scales. You are asked to indicate your own personal ideas about each subject as it is taught to CSE/O level using the given scales.

For example, if you think that a subject is extremely practical, you would place a tick on the extreme left-hand line.

Practical ✓ : ___ : ___ : ___ : ___ : ___ : ___ Theoretical
extremely quite slightly neutral slightly quite extremely

If you believe that a subject is extremely theoretical, you would put a tick on the extreme right-hand line.

Practical ___ : ___ : ___ : ___ : ___ : ___ : ✓ Theoretical
extremely quite slightly neutral slightly quite extremely

If your ideas about a subject are not so definite, a tick should be placed in the most appropriate intermediate position.

If you consider (a)a subject is neutral on the scale, or (b)both ends of the scale are equally associated with the subject, or (c)the scale is completely irrelevant to the subject, then you should place a tick on the middle line.

Practical ___ : ___ : ___ : ✓ : ___ : ___ : ___ Theoretical
extremely quite slightly neutral slightly quite extremely

It is your immediate response to each item that is required, even if some of the scales may seem inappropriate to describe school subjects. Therefore trust your first impressions and work quickly.

Do not attempt to compare your responses for the different subjects. Each tick should represent a separate and independent judgement. Additionally, do not confer with your colleagues until you have completed this questionnaire.

PLEASE BE SURE YOU TICK EVERY SCALE FOR EACH SUBJECT.

A3

FRENCH

Theoretical	__ :	__ :	__ :	__ :	__ :	__ :	__	Practical
Numerical	__ :	__ :	__ :	__ :	__ :	__ :	__	Verbal
Science	__ :	__ :	__ :	__ :	__ :	__ :	__	Arts
Logical	__ :	__ :	__ :	__ :	__ :	__ :	__	Intuitive
Feminine	__ :	__ :	__ :	__ :	__ :	__ :	__	Masculine
Factual	__ :	__ :	__ :	__ :	__ :	__ :	__	Opinionative
Creative	__ :	__ :	__ :	__ :	__ :	__ :	__	Routine
Simple	__ :	__ :	__ :	__ :	__ :	__ :	__	Complex
Important	__ :	__ :	__ :	__ :	__ :	__ :	__	Unimportant

CHEMISTRY

Theoretical	__ :	__ :	__ :	__ :	__ :	__ :	__	Practical
Numerical	__ :	__ :	__ :	__ :	__ :	__ :	__	Verbal
Science	__ :	__ :	__ :	__ :	__ :	__ :	__	Arts
Logical	__ :	__ :	__ :	__ :	__ :	__ :	__	Intuitive
Feminine	__ :	__ :	__ :	__ :	__ :	__ :	__	Masculine
Factual	__ :	__ :	__ :	__ :	__ :	__ :	__	Opinionative
Creative	__ :	__ :	__ :	__ :	__ :	__ :	__	Routine
Simple	__ :	__ :	__ :	__ :	__ :	__ :	__	Complex
Important	__ :	__ :	__ :	__ :	__ :	__ :	__	Unimportant

WOODWORK

Theoretical	__ :	__ :	__ :	__ :	__ :	__ :	__	Practical
Numerical	__ :	__ :	__ :	__ :	__ :	__ :	__	Verbal
Science	__ :	__ :	__ :	__ :	__ :	__ :	__	Arts
Logical	__ :	__ :	__ :	__ :	__ :	__ :	__	Intuitive
Feminine	__ :	__ :	__ :	__ :	__ :	__ :	__	Masculine
Factual	__ :	__ :	__ :	__ :	__ :	__ :	__	Opinionative
Creative	__ :	__ :	__ :	__ :	__ :	__ :	__	Routine
Simple	__ :	__ :	__ :	__ :	__ :	__ :	__	Complex
Important	__ :	__ :	__ :	__ :	__ :	__ :	__	Unimportant

MATHEMATICS

Theoretical	__ :	__ :	__ :	__ :	__ :	__ :	__	Practical
Numerical	__ :	__ :	__ :	__ :	__ :	__ :	__	Verbal
Science	__ :	__ :	__ :	__ :	__ :	__ :	__	Arts
Logical	__ :	__ :	__ :	__ :	__ :	__ :	__	Intuitive
Feminine	__ :	__ :	__ :	__ :	__ :	__ :	__	Masculine
Factual	__ :	__ :	__ :	__ :	__ :	__ :	__	Opinionative
Creative	__ :	__ :	__ :	__ :	__ :	__ :	__	Routine
Simple	__ :	__ :	__ :	__ :	__ :	__ :	__	Complex
Important	__ :	__ :	__ :	__ :	__ :	__ :	__	Unimportant

PHYSICS

Theoretical	___ : ___ : ___ : ___ : ___ : ___ : ___	Practical
Numerical	___ : ___ : ___ : ___ : ___ : ___ : ___	Verbal
Science	___ : ___ : ___ : ___ : ___ : ___ : ___	Arts
Logical	___ : ___ : ___ : ___ : ___ : ___ : ___	Intuitive
Feminine	___ : ___ : ___ : ___ : ___ : ___ : ___	Masculine
Factual	___ : ___ : ___ : ___ : ___ : ___ : ___	Opinionative
Creative	___ : ___ : ___ : ___ : ___ : ___ : ___	Routine
Simple	___ : ___ : ___ : ___ : ___ : ___ : ___	Complex
Important	___ : ___ : ___ : ___ : ___ : ___ : ___	Unimportant

HOME ECONOMICS

Theoretical	___ : ___ : ___ : ___ : ___ : ___ : ___	Practical
Numerical	___ : ___ : ___ : ___ : ___ : ___ : ___	Verbal
Science	___ : ___ : ___ : ___ : ___ : ___ : ___	Arts
Logical	___ : ___ : ___ : ___ : ___ : ___ : ___	Intuitive
Feminine	___ : ___ : ___ : ___ : ___ : ___ : ___	Masculine
Factual	___ : ___ : ___ : ___ : ___ : ___ : ___	Opinionative
Creative	___ : ___ : ___ : ___ : ___ : ___ : ___	Routine
Simple	___ : ___ : ___ : ___ : ___ : ___ : ___	Complex
Important	___ : ___ : ___ : ___ : ___ : ___ : ___	Unimportant

BIOLOGY

Theoretical	___ : ___ : ___ : ___ : ___ : ___ : ___	Practical
Numerical	___ : ___ : ___ : ___ : ___ : ___ : ___	Verbal
Science	___ : ___ : ___ : ___ : ___ : ___ : ___	Arts
Logical	___ : ___ : ___ : ___ : ___ : ___ : ___	Intuitive
Feminine	___ : ___ : ___ : ___ : ___ : ___ : ___	Masculine
Factual	___ : ___ : ___ : ___ : ___ : ___ : ___	Opinionative
Creative	___ : ___ : ___ : ___ : ___ : ___ : ___	Routine
Simple	___ : ___ : ___ : ___ : ___ : ___ : ___	Complex
Important	___ : ___ : ___ : ___ : ___ : ___ : ___	Unimportant

HISTORY

Theoretical	___ : ___ : ___ : ___ : ___ : ___ : ___	Practical
Numerical	___ : ___ : ___ : ___ : ___ : ___ : ___	Verbal
Science	___ : ___ : ___ : ___ : ___ : ___ : ___	Arts
Logical	___ : ___ : ___ : ___ : ___ : ___ : ___	Intuitive
Feminine	___ : ___ : ___ : ___ : ___ : ___ : ___	Masculine
Factual	___ : ___ : ___ : ___ : ___ : ___ : ___	Opinionative
Creative	___ : ___ : ___ : ___ : ___ : ___ : ___	Routine
Simple	___ : ___ : ___ : ___ : ___ : ___ : ___	Complex
Important	___ : ___ : ___ : ___ : ___ : ___ : ___	Unimportant

Finally, I would like to ask what characteristics you believe make a subject attractive to 14 year old pupils. Using the same 7-point scales as before, place ticks to show the extent to which pupils prefer the different characteristics.

A typical 14 year old <u>girl</u> prefers subjects which are:

Theoretical	__ : __ : __ : __ : __ : __ : __	Practical
Numerical	__ : __ : __ : __ : __ : __ : __	Verbal
Science	__ : __ : __ : __ : __ : __ : __	Arts
Logical	__ : __ : __ : __ : __ : __ : __	Intuitive
Feminine	__ : __ : __ : __ : __ : __ : __	Masculine
Factual	__ : __ : __ : __ : __ : __ : __	Opinionative
Creative	__ : __ : __ : __ : __ : __ : __	Routine
Simple	__ : __ : __ : __ : __ : __ : __	Complex
Important	__ : __ : __ : __ : __ : __ : __	Unimportant

A typical 14 year old <u>boy</u> prefers subjects which are:

Theoretical	__ : __ : __ : __ : __ : __ : __	Practical
Numerical	__ : __ : __ : __ : __ : __ : __	Verbal
Science	__ : __ : __ : __ : __ : __ : __	Arts
Logical	__ : __ : __ : __ : __ : __ : __	Intuitive
Feminine	__ : __ : __ : __ : __ : __ : __	Masculine
Factual	__ : __ : __ : __ : __ : __ : __	Opinionative
Creative	__ : __ : __ : __ : __ : __ : __	Routine
Simple	__ : __ : __ : __ : __ : __ : __	Complex
Important	__ : __ : __ : __ : __ : __ : __	Unimportant

Thank you for your cooperation and assistance.
If you would care to add any comments regarding this survey, I would welcome them.

Margaret Spear
The Open University

APPENDIX 4.7

THE COSS QUESTIONNAIRE (For science teachers)

CHARACTERISTICS OF SCHOOL SUBJECTS

The purpose of this survey is to investigate your views about several school subjects as they are taught in secondary schools up to CSE/O level standard.

PLEASE SUPPLY

(For statistical purposes only)

Sex: Female ☐ Male ☐

Age: Under 30 ☐ 40 - 49 ☐

30 - 39 ☐ 50 and over ☐

Your principal teaching subject ____________________

Subsidiary teaching subject(s) ____________________

DIRECTIONS

On the following pages you will find a selection of school subjects and beneath each a set of descriptive scales. You are asked to indicate your own personal ideas about each subject as it is taught to CSE/O level using the given scales.

For example, if you think that a subject is extremely practical, you would place a tick on the extreme left-hand line.

Practical ✓ : ___ : ___ : ___ : ___ : ___ : ___ Theoretical
extremely quite slightly neutral slightly quite extremely

If you believe that a subject is extremely theoretical, you would put a tick on the extreme right-hand line.

Practical ___ : ___ : ___ : ___ : ___ : ___ : ✓ Theoretical
extremely quite slightly neutral slightly quite extremely

If your ideas about a subject are not so definite, a tick should be placed in the most appropriate intermediate position.

If you consider (a)a subject is neutral on the scale, or (b)both ends of the scale are equally associated with the subject, or (c)the scale is completely irrelevant to the subject, then you should place a tick on the middle line.

Practical ___ : ___ : ___ : ✓ : ___ : ___ : ___ Theoretical
extremely quite slightly neutral slightly quite extremely

It is your immediate response to each item that is required, even if some of the scales may seem inappropriate to describe school subjects. Therefore trust your first impressions and work quickly.

Do not attempt to compare your responses for the different subjects. Each tick should represent a separate and independent judgement. Additionally, do not confer with your colleagues until you have completed this questionnaire.

PLEASE BE SURE YOU TICK EVERY SCALE FOR EACH SUBJECT.

CHEMISTRY

Theoretical	___	___	___	___	___	___	___	Practical
Numerical	___	___	___	___	___	___	___	Verbal
Science	___	___	___	___	___	___	___	Arts
Logical	___	___	___	___	___	___	___	Intuitive
Feminine	___	___	___	___	___	___	___	Masculine
Factual	___	___	___	___	___	___	___	Opinionative
Creative	___	___	___	___	___	___	___	Routine
Simple	___	___	___	___	___	___	___	Complex
Important	___	___	___	___	___	___	___	Unimportant

MATHEMATICS

Theoretical	___	___	___	___	___	___	___	Practical
Numerical	___	___	___	___	___	___	___	Verbal
Science	___	___	___	___	___	___	___	Arts
Logical	___	___	___	___	___	___	___	Intuitive
Feminine	___	___	___	___	___	___	___	Masculine
Factual	___	___	___	___	___	___	___	Opinionative
Creative	___	___	___	___	___	___	___	Routine
Simple	___	___	___	___	___	___	___	Complex
Important	___	___	___	___	___	___	___	Unimportant

PHYSICS

Theoretical	___	___	___	___	___	___	___	Practical
Numerical	___	___	___	___	___	___	___	Verbal
Science	___	___	___	___	___	___	___	Arts
Logical	___	___	___	___	___	___	___	Intuitive
Feminine	___	___	___	___	___	___	___	Masculine
Factual	___	___	___	___	___	___	___	Opinionative
Creative	___	___	___	___	___	___	___	Routine
Simple	___	___	___	___	___	___	___	Complex
Important	___	___	___	___	___	___	___	Unimportant

BIOLOGY

Theoretical	___	___	___	___	___	___	___	Practical
Numerical	___	___	___	___	___	___	___	Verbal
Science	___	___	___	___	___	___	___	Arts
Logical	___	___	___	___	___	___	___	Intuitive
Feminine	___	___	___	___	___	___	___	Masculine
Factual	___	___	___	___	___	___	___	Opinionative
Creative	___	___	___	___	___	___	___	Routine
Simple	___	___	___	___	___	___	___	Complex
Important	___	___	___	___	___	___	___	Unimportant

Next, I would like to ask what characteristics you believe make a subject attractive to 14 year old pupils. Using the same 7-point scales as before, place ticks to show the extent to which pupils prefer the different characteristics.

A typical 14 year old girl prefers subjects which are:

Theoretical	__ : __ : __ : __ : __ : __ : __	Practical
Numerical	__ : __ : __ : __ : __ : __ : __	Verbal
Science	__ : __ : __ : __ : __ : __ : __	Arts
Logical	__ : __ : __ : __ : __ : __ : __	Intuitive
Feminine	__ : __ : __ : __ : __ : __ : __	Masculine
Factual	__ : __ : __ : __ : __ : __ : __	Opinionative
Creative	__ : __ : __ : __ : __ : __ : __	Routine
Simple	__ : __ : __ : __ : __ : __ : __	Complex
Important	__ : __ : __ : __ : __ : __ : __	Unimportant

A typical 14 year old boy prefers subjects which are:

Theoretical	__ : __ : __ : __ : __ : __ : __	Practical
Numerical	__ : __ : __ : __ : __ : __ : __	Verbal
Science	__ : __ : __ : __ : __ : __ : __	Arts
Logical	__ : __ : __ : __ : __ : __ : __	Intuitive
Feminine	__ : __ : __ : __ : __ : __ : __	Masculine
Factual	__ : __ : __ : __ : __ : __ : __	Opinionative
Creative	__ : __ : __ : __ : __ : __ : __	Routine
Simple	__ : __ : __ : __ : __ : __ : __	Complex
Important	__ : __ : __ : __ : __ : __ : __	Unimportant

SUBJECT CHOICE

Finally, a couple of questions about the subjects that pupils choose to study for their 16+ examinations.

How important do you think that CSE/O level qualifications in the following subject areas will be to girls in their future lives?

GIRLS	Very important	Fairly important	Not very important	Not at all important
Creative Arts				
Languages				
Humanities				
Science				
Technical Subjects				
Home Economics				
Commercial/Business Studies				

How important do you think that CSE/O level qualifications in the following subject areas will be to boys in their future lives?

BOYS	Very important	Fairly important	Not very important	Not at all important
Creative Arts				
Languages				
Humanities				
Science				
Technical Subjects				
Home Economics				
Commercial/Business Studies				

Thank you for your cooperation.
If you would care to add any comments regarding the aims and/or design of this survey, I would welcome them.

Margaret Spear

Table A4.8/1 Sex of respondents

Sex	No. of teachers	Percentage (adjusted)	National percentage
Male	236	66.7	66
Female	118	33.3	34
Unknown	3		

Source: HMI (1979)

Table A4.8/2 Principal teaching subject of respondents

Subject	No. of teachers	Percentage (adjusted)	National percentage
Physics	85	24.4	26.1
Chemistry	110	31.6	23.2
Biology	115	33.1	33.3
Integrated Science	38	10.9	17.4
Other	4		
Unknown	5		

Source: HMI (1979)

Table A4.8/3 Sex ratio within different teaching subjects

Subject	No. men	No. women	% men	% men nationally	χ^2	p
Physics	71	14	83.5	88	1.61	0.30
Chemistry	84	24	77.8	81	1.74	0.50
Biology	56	58	49.1	46	0.46	0.50
Integrated Science	20	18	52.6	68	4.06	0.05

Source: HMI (1979)

Table A4.8/4 Age distribution of respondents

Age range (years)	No. of teachers	Percentage (adjusted)	National percentage
Under 30	81	23.9	43.2
30 - 39	158	46.6	30.1
40 - 49	65	19.2	16.4
50 and over	35	10.3	10.3
Unknown	18		

Source: Chapman (1980)

APPENDIX 4.8

SUPPLEMENTARY INFORMATION ABOUT THE BIAS AND STOSS RESPONDENTS

The sex ratio of the BIAS and STOSS sample is compared in Table A4.8/1 with that obtained for science teachers in a large sample of maintained secondary schools visited by H. M. inspectors (HMI, 1979). Calculation of χ^2 indicates that the sex ratio of the BIAS and STOSS sample is representative of the estimated national ratio ($\chi^2 = 0.073$, $p = 0.80$). Table A4.8/2 records proportions of BIAS and STOSS respondents teaching the different science subjects compared with proportions found in the HMI sample. The χ^2 value is highly significant indicating that the sample is biased with respect to the subjects taught by the respondents. This probably occurred because greater numbers of chemistry teachers were willing to mark the chemistry slanted experimental write-ups. Table A4.8/3 records the sex ratio within each group of subject teachers, and shows that the sex ratio is representative of that found by the H. M. inspectors in every science subject, with the exception of the integrated science teachers.

The age distribution of the BIAS and STOSS sample can be seen in Table A4.8/4, together with national figures supplied by Chapman (1980). It can be seen that the sample is under-represented as regards the younger, less experienced science teacher. This point is further brought out in Table A4.8/5, which shows the respondents' teaching experience. In the absence of data referring specifically to the teaching experience of all science teachers, the national figures appearing in Table A4.8/5 refer to teachers of all subjects, and therefore may not accurately reflect the teaching experience of science teachers.

Table A4.8/6 records the qualifications of the BIAS and STOSS sample. No national data are available for comparison, apart from the observation by H. M. inspectors (HMI, 1979) that 70% of their sample of science teachers were graduates, of whom 80% had followed a course of

Table A4.8/5 Teaching experience of respondents

Teaching experience (years)	No. of teachers	Percentage (adjusted)	National percentage
Less than 2	25	7.4	19
2 - 5	61	18.0	20
5 - 10	94	27.7	22
10 - 20	115	33.9	23
Over 20	44	13.0	16
Unknown	18		

Source: DES, Statistical Bulletin, 6/80

Table A4.8/6 Qualifications of respondents

Qualification	No. of teachers	Percentage (adjusted)
Certificate of Education	50	14.8
BEd	25	7.4
Untrained graduate	48	14.2
Trained graduate	215	63.6
Unknown	19	

Table A4.8/7 Qualifications of different subject teachers

Subject	Number trained graduates	No. with other qualifications	% trained graduates	% trained graduates nationally	χ^2	p
Physics	52	27	65.8	60	1.12	0.3
Chemistry	77	29	72.6	66	2.06	0.2
Biology	66	43	60.6	57	0.57	0.5
Integrated Science	17	19	47.2	36	1.64	0.2

Table A4.8/8 Teaching experience of trained gradiates

Teaching experience (years)	No. of teachers	Percentage	National percentage
Less than 10	119	55.3	59
More than 10	96	44.7	41

Source: DES, Statistical Bulletin, 6/80

professional training. 85% of the BIAS and STOSS sample were graduates, and 83% of those graduates were trained to teach. Table A4.8/7 records the number of trained graduates within each group of subject teachers, and shows that the ratio of trained graduates to teachers with other qualifications is representative of that found by the H. M. inspectors in all the science subjects. Table A4.8/8 shows that when the variable teaching experience is dichotomised into less than or more than 10 years, then the teaching experience of the trained graduates in the sample is representative of national figures referring to teachers of all subjects ($\chi^2 = 1.18$, $p = 0.3$).

Sixty two (18%) of the teachers in the BIAS and STOSS sample were overall head of science in their respective schools. Of these heads of science, 82% were men and 87% were graduates. Comparable figures obtained in the HMI survey were 84% and 76%. The BIAS and STOSS sample of heads of science, besides being well qualified, were highly experienced teachers, 89% of them having taught for more than 10 years.

APPENDIX 4.9

THE SCHOOLS INVOLVED IN THE STUDY

Details of the schools which participated in the research appear in Table A4.9/1. Reading this table in conjunction with Table A4.12/1 in Appendix 4.12 shows the characteristics of the schools involved in all the different stages and aspects of the investigation.

Table A4.9/1 Details of the participating schools

Study (Sample code)	No. of schools involved	T Type of schools	Sex of schools	Location of schools
A	5	Comprehensives Secondary modern	Mixed	Home counties
A*	4	Comprehensives	Mixed	Home counties
C	5	Independent	Mixed (2) Boys' (2) Girls'(1)	S.E. England East Anglia
D	4	Comprehensives	Mixed	E. & N.W. of England
E	2	Comprehensives	Mixed	Home counties, Middlesex
F	6	Comprehensives	Mixed	S.E. & N. of England
H	3	Comprehensives	Mixed	S. & S.E. of England Midlands
I	9	Comprehensives	Mixed	S.W. & N. of England Home counties
I*	7	Comprehensives	Mixed	S.W. & N.W. England Midlands
J	1	Comprehensive	Mixed	Midlands
K	6	Comprehensives	Mixed	S.W. & S.E. England
L	6	Comprehensives	Mixed	Home counties East Anglia
M	25	Comprehensives	Mixed	Throughout England
N	13	Middle	Mixed	Throughout England
O	18	Primary	Mixed	Throughout England
P	86	Comprehensives Grammar Secondary moderns	Mixed Boy' Girls'	Throughout England
TSCH	9	Comprehensives	Mixed	Home counties
INT	5	Comprehensives Secondary modern	Mixed	Home counties

Table A4.10/1 Type of school in which respondents taught (BIAS and STOSS sample)

Type of school	No. of teachers	Percentage (adjusted)	National percentage
Secondary modern	27	7.6	6.1
Grammar	26	7.3	3.7
Comprehensive	294	82.8	89.0
Other	8	2.3	1.2
Unknown	2		

Table A4.10/2 Sex of school in which respondents taught (BIAS and STOSS sample)

Sex of school	No. of teachers	Percentage (adjusted)	National percentage
Mixed	277	78.0	80.0
Boys'	21	5.9	9.7
Girls'	57	16.1	10.3
Unknown	2		

Table A4.10/3 Type and sex of school in which respondents taught (BIAS and STOSS sample)

Type and sex of school	No. of teachers	Percentage	National percentage
Comprehensive			
Mixed	247	69.6	71.6
Single sex	47	13.2	12.2
Other			
Mixed	30	8.5	8.4
Single sex	31	8.7	7.8

Table A4.10/4 Size of school in which respondents taught (BIAS and STOSS sample)

Size of school (No. of pupils)	No. of teachers	Percentage (adjusted)	National percentage
Under 600	16	4.5	21.5
601 - 1000	164	46.3	43.3
1001 and over	174	49.2	35.2
Unknown	3		

APPENDIX 4.10

SUPPLEMENTARY INFORMATION ABOUT THE SCHOOLS THAT RETURNED BIAS, STOSS AND COSS QUESTIONNAIRES

The type of school in which BIAS and STOSS respondents taught is recorded in Table A4.10/1. The percentage of teachers in the different types of all maintained secondary school in England and Wales is included for comparison. Table A4.10/2 shows the number of respondents teaching in mixed and single sex schools. The percentages can be compared with the corresponding percentages calculated for all maintained secondary schools in England and Wales. Combining the information contained in Tables A4.10/1 and A4.10/2, a school type x school sex breakdown for the BIAS and STOSS sample can be calculated (Table A4.10/3). Corresponding national percentages, calculated from DES Statistics of Schools 1981, are also included. It can be seen that the majority of state secondary schools are mixed comprehensives (71.6%), and that the sample accurately reflects this situation, since 69.6% of the respondents taught in mixed comprehensives. χ^2 was calculated to determine the goodness of fit between the sample figures and national figures. A value of 0.91 ($p=0.90$) was obtained, which suggests that the null hypothesis should not be rejected, i.e. that the sample is representative of the population, with respect to major divisions between school type and school sex. Furthermore, the fact that 95.5% of the respondents were teaching in schools which had not undergone reorganisation within the previous two years, indicates that the sample schools were stable representatives of their type.

Table A4.10/4 shows the size of the schools in which the respondents taught. The figures do not compare well with national statistics. School size distributions are probably linked with age range distributions, which are also poor. DES statistics indicate that 61.9% of secondary schools in England and Wales teach pupils up to A level. In

Table A4.10/5 Age range of school in which respondents taught (Primary COSS sample)

Type of school	Age range	No. of teachers
First	5 - 9	11
Junior with infants	5 - 11	40
Junior	7 - 11	51

Table A4.10/6 Catchment area of schools in which respondents taught (BIAS and STOSS sample)

Catchment area	No. of teachers	Percentage (adjusted)
Large city - inner	35	11
Large city - suburban	98	31
Large town	130	41
Rural	54	17
Unknown	40	

Table A4.10/7 Background of pupils taught by respondents (BIAS and STOSS sample)

Background of pupils	No. of teachers	Percentage (adjusted)
Prosperous	79	24.5
Average	135	41.8
Disadvantaged	16	4.9
Mixed	93	28.8
Unknown	34	

Source (Tables A4.10/1 to A4.10/4): DES Statistics of Schools 1981 the BIAS and STOSS sample, 79.8% of respondents taught in schools which took pupils up to 18, whereas the figure for the secondary school COSS sample was only 35.2%. Age range distributions for the primary school COSS sample are recorded in Table A4.10/5.

Information was also sought concerning the location of schools answering the BIAS and STOSS questionnaires, and the economic backgound of their pupils. Table A4.10/6 shows the number of respondents teaching in schools with different catchment areas, and Table A4.10/7 refers to the background of the pupils taught by the respondents. These two tables indicate that the respondents taught in schools which varied widely as regards their location and the type of pupil that they catered for. No national figures are available for comparison.

Table A4.11/1 Details of the teacher samples

Sample code	Total no.	No. of males	No. of females	No. of non-science teachers	Science teachers: Total no.	No. of physics teachers	No. of chemistry teachers	No. of biology teachers	Combined science teachers
A, A*	35	27	8	0	35	7	8	10	10
C	35	26	9	27	8	2	3	3	0
D	49	18	31	49	0				
E	25	14	11	21	4	1	1	2	0
F	77	42	35	53	24	8	7	5	4
H	67	30	37	55	12	3	3	4	2
I	36	27	9	0	36	6	12	12	6
I*	35	-	-	0	35	9	10	14	2
J	36	16	20	32	4	2	0	2	0
K	45	35	10	0	45	13	11	17	4
L	49	31	18	0	49	17	13	14	5
M	290	(167)	(110)	237	53	-	-	-	-
N	120	(47)	(68)	106	14	-	-	-	-

APPENDIX 4.11

THE TEACHERS INVOLVED IN THE STUDY

Table A4.11/1 contd.

Sample code	Total no.	No. of males	No. of females	No. of non-science teachers	Total no.	No. of physics teachers	No. of chemistry teachers	No. of biology teachers	Combined science teachers
O	102	(29)	(62)	102	0				
P	357	(236)	(118)	0	357	(85)	(110)	(115)	(42)
TSCH	67	(44)	(22)	0	67	18	20	20	9
INT	23	15	8	0	23	8	4	8	3

Note

( ) Due to missing data, not all respondents are accounted for.

\- Information not available

Table A4.11/2 Details of the PGCE student sample

Sample code	No. of males	No. of females	No. of non-science students	Science students: Total no.	No. of physics students	No. of chemistry students	No. of biology students
ST(BIAS)	14	7	0	21	2	16	3
ST(STOSS)	18	10	0	28	12	4	12
ST(COSS)	18	17	20	15	-	-	-

APPENDIX 4.12

THE INVOLVEMENT OF THE SAMPLES IN DATA COLLECTION

This appendix consists of four tables, which summarize the involvement of the samples in the development of the scales and their eventual use. The first table lists all the different teacher samples that were involved in the study, and shows the different scales that each sample was asked to complete. The second table indicates the date that each sample was approached, and the manner in which questionnaires were distributed to the respondents. The third table describes the collection of data from the PGCE students for the reliability studies. The last table details the samples used in the development and final use of each scale. Besides showing the number of different stages that each scale passed through, the table also contains a complete list of all the scales that were used in the research, including those that were only designed to collect background data required for the further development of the final measures. The speed at which scales were developed and piloted can be ascertained by reading Table A4.12/4 in conjunction with Table A4.12/2.

Table A4.12/1 Testing schedule (1) Scales

Sample code	Sample size	Respondents	Scales	Form
A&A*	35	Science teachers	Personal details	Pilot (1)
			Marking exercise	Prelim (1)
			Variables used in marking	Prelim
			Cards	Pilot (1)
			~~Masculinity Index~~	~~Pilot (1)~~
			Characteristics of science	Pilot (1)
			Reasons for choosing/dropping	Pilot (1)
			Reasons for success/failure	Pilot (1)
			Scientist stereotypes	Pilot (1)
			Written work of girls and boys	Prelim
B	20	WI members	Gender of word pairs	Prelim (1)
			Name preferences	Prelim (1)
C	35	Secondary teachers	School subject characteristics	Pilot (1)
			Preference for subject characteristics	Pilot (1)
			Opinions	Pilot (1)
D	49	Secondary teachers	Gender of word pairs	Prelim (2)
			Name preferences	Prelim (2)
			Written work of girls and boys	Final
E	25	Secondary teachers	Gender of word pairs	Prelim (3)
			Name preferences	Prelim (3)
			Written work of girls and boys	Final
F	77	Secondary teachers	School subject characteristics	Pilot (2)
			Preference for subject characteristics	Pilot (2)
			Opinions	Final
H	67	Secondary teachers	School subject characteristics	Pilot (3)
			Preference for subject characteristics	Pilot (3)
			Opinions	Final
			Importance of subjects	Pilot (1)
I	36	Science teachers	Personal details	Pilot (2)
			Marking exercise	Pilot
			Cards	Pilot (2)
			(Plus the 5 scales administered to sample I*)	

I*	35	Science teachers	Masculinity Index	Pilot
			Characteristics of science	Pilot (2)
			Reasons for choosing/dropping	Pilot (2)
			Reasons for success/failure	Pilot (2)
			Scientist stereotypes	Pilot (2)
J	36	Secondary teachers	Importance of subjects	Pilot (2)
			Adjective pairs	Prelim (1)
			Gender of word pairs	Prelim (4)
			Name preferences	Prelim (4)
			Written work of girls and boys	Final
K	45	Science teachers	Females' social roles	Pilot (1)
			Masculinity Index	Final
			Marking exercise	Prelim (2)
L	49	Science teachers	Females' social roles	Pilot (2)
			Masculinity Index	Final
			Adjective pairs	Prelim (2)
			Written work of girls and boys	Final
M	290	Secondary teachers	School subject characteristics	Final
			Preference for subject characteristics	Final
N	120	Middle school teachers	School subject characteristics	Final
			Preference for subject characteristics	Final
O	102	Primary teachers	School subject characteristics	Final
			Preference for subject characteristics	Final
P	357	Science teachers	Personal details	Final
			Marking exercise	Final
			Masculinity Index	Final
			Characteristics of science	Final
			Preference for subject characteristics	Final
			Reasons for choosing/dropping	Final
			Reasons for success/failure	Final
			Scientist stereotypes	Final
			Females' social roles	Final
TSCH	67	Science teachers	School subject characteristics	Final
			Importance of subjects	Final

Table A4.12/2 Testing schedule (2) Administration

Sample code	Sampling unit	Method of administration	Date of administration
A	Science dept.	Personal direct	Dec. 1980
A*	Science dept.	Via head of science	Dec. 1980
B	WI group	Via acquaintance	Oct. 1980
C	School	Via head teacher	Dec. 1980
D	School	Via assistant teacher	Dec. 1980/Jan. 1981
E	School	Via assistant teacher	Feb. 1981
F	School	Via head teacher	Feb. 1981
H	School	Via head teacher	March 1981
I	Science dept.	Via head of science	March 1981
I*	Science dept.	Via head of science	Oct. 1981
J	School	Via assistant teacher	May 1981
K	Science dept.	Via head of science	May/June 1981
L	Science dept.	Via head of science	Oct. 1981
M	School	Via head teacher	Oct. 1981
N	School (middle)	Via head teacher	Oct. 1981 Feb. 1982
O	School (primary)	Via head teacher	Oct. 1981 Feb./March 1982
P	Science dept.	Via head of science	Oct.-Dec. 1981 Feb./March 1982 May 1982
TSCH	Science dept.	Via head of science	June 1981
INT	Science dept.	Personal direct	June 1980

Table A4.12/3 Testing schedule (3) Reliability studies

Sample code	Sampling unit	Method of administration	Dates of completion
ST(BIAS)	UDE	Via science education lecturer	April & May 1982
ST(STOSS)	UDE	Via science education lecturer	April & May 1982
ST(COSS)	UDE	Via science education lecturer	April & May 1982

Table A4.12/4 Scales: Their development and use

Scale	Use†		
	Prelim	Pilot	Final form
Adjective pairs	J, L		
Cards		A&A*, I	
Characteristics of science		A&A*, I&I*	P
Females' social roles		K, L	P
Gender of word pairs	B, D, E, J		
Importance of subjects		H, J	TSCH
Marking exercise	A&A*, K	I	P
Masculinity Index		C, F	K, L, P
Name preferences	B, D, E, J		
Opinions		C	F, H
Personal details		A&A*, I	P
Preference for subject characteristics		C, F, H	M, N, O, P
Reasons for choosing/dropping		A&A*, I&I*	P
Reasons for success/failure		A&A*, I&I*	P
School details		A&A*, I&I*	P, O
School subject characteristics		C, F, H	M, N, O
Scientist stereotypes		A&A*, I&I*	P
Variables used in marking	A&A*		
Written work of girls and boys	A&A*		D, E, J, L

† The letters in the table refer to sample codes and identify the respondents

APPENDIX 4.13

PRINCIPAL QUESTIONNAIRE RETURNS

The BIAS, STOSS and COSS questionnaires all appeared in more than one format. Roughly equal numbers of the different formats of each questionnaire were returned. Full details of the returns of each questionnaire are shown in Table A4.13/1 to A4.13/4.

Tables A4.13/1 and A4.13/2 refer to the BIAS questionnaires that were returned. Table A4.13/1 shows the number of sample pairs (write-up plus essay) of each quality that had been marked in association with a boy's name, and the number that had been marked when linked with a girl's name. The percentage split between boys' work and girls' work is shown in brackets. Table A4.13/2 records the number and percentage of returned questionnaires which contained the high standard work in the first, second and third position. Corresponding figures are also given for the other two standards of work.

Tables A4.13/3 and A4.13/4 record the returns of the two formats of the STOSS and COSS questionnaires. Table A4.13/3 shows the number and percentage of teachers who returned the 'girl' and 'boy' format of the STOSS questionnaire. Table A4.13/4 shows the number and percentage of teachers who returned the two different formats of the COSS questionnaire. The formats differed in the range of subjects presented for rating. Additional variation arose in the COSS questionnaire since the subjects were presented in different orders. Table A4.13/4 also shows the number and percentage of teachers who completed each of the eight different subject arrangements.

Table A4.13/1 BIAS returns (1) Pupil sex frequencies

Sample pair	Pupil sex	
	Boy	Girl
Good	146 (43.1%)	193 (56.9%)
Average	200 (59.0%)	139 (41.0%)
Poor	162 (47.8%)	177 (52.2%)

Table A4.13/2 BIAS returns (2) Sample frequencies in each position

Sample pair	Position in booklet		
	1st	2nd	3rd
Good	106 (31.3%)	113 (33.3%)	120 (35.4%)
Average	121 (35.7%)	107 (31.6%)	111 (32.7%)
Poor	112 (33.0%)	120 (33.6%)	107 (30.0%)

Table A4.13/3 STOSS returns

Format	Number	Percentage
Boy	73	44.5
Girl	91	55.5

Table A4.13/4 COSS returns

(A) COLLECTION OF SUBJECTS

Format	Number	Percentage
A	256	50
B	256	50

Note A = physics, home economics, biology, history, French, chemistry, woodwork, maths

B = geography, physics, art, biology, chemistry, English (language), maths, technical drawing

(B) POSITION OF SUBJECTS

Arrangement	Number	Percentage
1	62	12.1
2	57	11.1
3	76	14.8
4	66	12.9
5	51	10.0
6	67	13.1
7	67	13.1
8	66	12.9

APPENDIX 5.1

INTERVIEW QUESTIONS

1. What subjects do you teach?
2. If school subjects were divided into science and non-science, what label would you use for the non-science group?
3. Which subjects form the (arts) group?
4. Which subjects form the science group?
5. If you were asked to divide science subjects into two broad groups, how would you split them?
6. How do science subjects differ from (arts) subjects?
7. Why do more girls than boys study biological subjects?
8. Why do more boys than girls study physical science subjects?
9. How does the work and behaviour of girls differ from that of boys in your subject?
10. What factors contribute to pupils' success in your subject area?
11. What factors cause pupils to fail in your subject?
12. Why do most girls drop physical science when choosing their subject options at 13+ or thereabouts?
13. How could the physical sciences be made more attractive for girls?
14. Why are the physical sciences often described as boys' subjects?
15. What qualities are required in order to become a successful research scientist?
16. What did you understand by the term 'research scientist' in the previous question?
17. How are scientists commonly portrayed?
18. Are the terms 'biologist' and 'physicist' meaningful to you?

APPENDIX 5.2

DETAILED SUMMARY OF THE INTERVIEW DATA

The tables appearing below present a brief synopsis of all the data contained in the interviews that were relevant to the six broad topics under investigation, i.e. school subjects, pupils' science choices, the masculine image of science, causes of success and failure at science, differences between girls and boys, and scientists. All of the data referring to each topic is gathered together and recorded in a single table, with the exception of the topic Pupils' Science Choices. This topic is covered by four separate tables. An additional table presents data on a closely related topic, that of encouraging more girls into physical science.

In each table, the individual points mentioned by the respondents have been organised into categories. The categories are ranked according to the total number of mentions made. Similarly, within each category the individual points are mostly ranked according to the frequency with which they were mentioned. In both cases the total number of mentions made is recorded in brackets. The total number of mentions made within each category do not necessarily indicate the number of teachers who referred to the category, since some teachers mentioned several points within a single category, and each separate point has been counted. In contrast, the figures beside the individual points do show the number of teachers who mentioned that point. Finally, it should be noted that although the sample size was 23, not all of the respondents discussed all of the topics. Thus, effective sample size was highly variable.

Table A5.2/1 Ways in which science subjects differ from arts subjects

1. Practical - theoretical (17)	
Practical (10)	Theoretical (1)
Active (2)	Mental (2)
Observations (1)	Constraints (1)
2. Factual - opinionative (13)	
Factual (7)	Opinions (4)
Preciseness (1)	
Explains (1)	
3. Analytic - synthetic (9)	
Analytical (4)	Logical (4)
	Deductive (1)
4. Structured - unstructured (7)	
Structured (5)	Open-ended (1)
	Diffuse (1)
Creative - derivative (7)	
Creative (5)	
Discovery (2)	
6. Objective - subjective (6)	
Objective (3)	Feelings (1)
Ideas (1)	Subjective (1)
7. Numerical - verbal (5)	
Mathematical (2)	Vocabulary (1)
Measurement (1)	Verbal (1)
8. Thing oriented - person oriented (3)	
Thing oriented (2)	Person oriented (1)
9. Difficult - easy (1)	
Difficult (1)	

Table A5.2/2 Why girls choose biology

1. Social factors (16)	
Society (6)	Encouragement (2)
Tradition (4)	Expectation (1)
Toys (3)	
2. Affective factors (13)	
Interest (8)	Clean (1)
Feminine subject (4)	
Future family life (13)	
Motherhood (5)	Home (2)
Children (4)	Health (2)
4. Subject characteristics (12)	
Descriptive (2)	Rote learning (1)
Little practical (2)	Structured (1)
Graphical aspects (2)	Not technical (1)
Less maths (2)	Easier (1)
5. Future working life (5)	
Careers (5)	
6. Teaching group composition (4)	
Majority (2)	With friends (2)
7. Influential people (3)	
Parents (2)	Teacher expectation (1)
8. School organisation (1)	
Modelling (1)	

Table A5.2/3 Why boys reject biology

1. Affective factors (3)	
Negativity (2)	No interest (1)
2. Future working life (2)	
No careers (2)	
Social factors (2)	
Tradition (1)	Sex roles (1)
4. Subject characteristics (1)	
Structured (1)	

Table A5.2/4 Why boys choose physics

1. Social factors (23)
    - Society (6)
    - Expectations (5)
    - Tradition (4)
    - Encouragement (3)
    - Toys (2)
    - Sex role stereotyping (2)
    - Modelling (1)
2. Subject characteristics (14)
    - Practical (5)
    - Mechanical (4)
    - Easy (2)
    - Less writing (1)
    - Logical (1)
    - Spatial abilities (1)
3. Future working life (11)
    - Careers (11)
4. Influential people (8)
    - Parents (5)
    - Teachers' attitudes (2)
    - Peer pressure (1)
5. Affective factors (6)
    - Masculine (4)
    - Interest (2)
6. Teaching group composition (3)
    - With friends (1)
    - Majority (1)
    - Teacher's sex (1)
7. School organisation (1)
    - Timetable (1)

Table A5.2/5 Why girls reject physics

1. Social factors (28)
    - Society (9)
    - No expectations (6)
    - Tradition (5)
    - Sex role stereotyping (4)
    - Prejudice (3)
    - Lack of models (1)
2. Affective factors (22)
    - Masculine (9)
    - No interest (7)
    - Dirty (3)
    - Dislike teacher (2)
    - Lack of confidence (1)
3. Subject characteristics (17)
    - Mechanical (6)
    - Mathematical (4)
    - Difficult (3)
    - Practical (2)
    - Lack of familiarity (1)
    - Spatial ability (1)
4. Influential people (11)
    - Parents (5)
    - Peer pressure (4)
    - Poor advice (1)
    - Teacher bias (1)
5. Future working life (6)
    - Unrelated to careers (6)

Teaching group composition (6)
- Minority (4)
- Rough boys (1)
- Teacher's sex (1)

7. School organisation (4)
    - Subject combination (2)
    - Subject competition (2)

Table A5.2/6 Ways of encouraging more girls into physical science

1. Future working life (6)	
Careers advice (6)	
Subject characteristics (6)	
Alter syllabus (4)	Feminine examples (1)
Minimise maths (1)	
3. Teaching group composition (5)	
Female teachers (3)	Girl groups (2)
4. Influential people (4)	
Enthusiastic staff (2)	Teacher encouragement (2)
5. Social factors (3)	
Change girls' attitudes(1)	Change society's attitudes (1)
Change schools' attitudes (1)	
6. School organisation (1)	
Compulsory science (1)	

Table A5.2/7 Reasons why science has a masculine image

1. Social factors (19)	
Tradition (11)	Toys (2)
Society (6)	
2. Man's role (9)	
Jobs (2)	Male domination (2)
Male scientists (2)	Power (1)
Man's role (2)	
3. Subject characteristics (5)	
Mathematical (2)	Dangers (1)
Mechanical (2)	
4. Influential people (4)	
Parental attitudes (2)	Peer attitudes (1)
Teachers' attitudes (1)	
5. Teaching group composition (3)	
Male pupils (2)	Male teachers (1)
6. School organisation (2)	
Boys' schools (2)	

Table A5.2/8 Causes of academic success in science

1. Approach to work (24)
   - Effort (8)
   - Motivation (3)
   - Goal (2)
   - Self-discipline (2)
   - Methodical (2)
   - Neatness (2)
   - Systematic (1)
   - Conscientiousness (1)
   - Obedience (1)
   - Concentration (1)
   - Curiosity (1)
2. Cognitive skills (15)
   - Knowledge (4)
   - Logic (3)
   - Application (3)
   - Maths (3)
   - Memory (1)
   - Observation (1)
3. Environmental factors (12)
   - Teacher (5)
   - Family support (4)
   - Facilitated success (3)
4. Inherent factors (11)
   - Ability (8)
   - Flair (3)
5. Affect (6)
   - Interest (5)
   - Attitude (1)
6. Behaviour (1)
   - Quiet (1)

Table A5.2/9 Causes of academic failure in science

1. Approach to work (18)
   - Lack of motivation (5)
   - Lack of perseverance (3)
   - Lack of effort (3)
   - Laziness (3)
   - Lack of patience (2)
   - Lack of curiosity (1)
   - Worry (1)
2. External factors (9)
   - Home factors (2)
   - Peer pressure (2)
   - Unavoidable failure (2)
   - Teacher (1)
   - Volume of work (1)
   - Vocabulary (1)
3. Cognitive skills (8)
   - Poor interpretation (2)
   - Poor English (2)
   - Poor maths (1)
   - Poor memory (1)
   - Rote learning (1)
   - Unfamiliarity (1)
4. Inherent factors (7)
   - Limited ability (7)
5. Affect (5)
   - Lack of interest (5)
6. Behaviour (1)
   - Poor behaviour (1)

Table A5.2/10 Differences between girls and boys

1. Work (29)
   - Neatness (13)
   - Results (7)
   - Verbal ability (3)
   - Written work (2)
   - Diagrams (1)
   - Relevance (1)
   - Confusion (1)
   - Speed (1)
2. Approach to work (24)
   - Effort (5)
   - Conscientiousness (4)
   - Motivation (3)
   - Curiosity (3)
   - Questioning (2)
   - Teacher influenced (2)
   - Confidence (2)
   - Tinker (1)
   - Concentration (1)
   - Sloth (1)
3. Affective differences (13)
   - Interest (6)
   - Attitude (4)
   - Appeal of mechanics (3)
4. Cognitive differences (12)
   - Ability (6)
   - Understanding (2)
   - Maths (2)
   - Logical (1)
   - Rote learning (1)
5. Behaviour (10)
   - Behaviour (5)
   - Chatter (4)
   - Active (1)
6. Teaching method preferences (7)
   - Practical (5)
   - Structured (2)
7. Physical differences (5)
   - Maturity (3)
   - Strength (2)

Table A5.2/11 Qualities required to become a research scientist

1. Personality characteristics (32)
   - Motivation (7)
   - Dedication (6)
   - Single-mindedness (4)
   - Perseverance (4)
   - Patience (4)
   - Determination (2)
   - Miserly (2)
   - Phlegmatic (1)
   - Humour (1)
   - Persuasive (1)
2. Inherent factors (14)
   - Intelligence (9)
   - Inspiration (5)
3. Cognitive skills (12)
   - Logical (6)
   - Enquiring mind (5)
   - Mathematical (1)
4. Work characteristics (12)
   - Work hard (4)
   - Plan ahead (2)
   - Practical (2)
   - Improvise (1)
   - Search (1)
   - Independent work (1)
   - Observation (1)
5. Affect (7)
   - Interest (4)
   - Enthusiasm (3)

APPENDIX 5.3

WAYS OF ENCOURAGING MORE GIRLS INTO PHYSICAL SCIENCE

During the exploratory interviews, the teachers were asked why it is that most girls drop physical science when choosing their subject options at 13+ or thereabouts. The following question asked for possible ways of combatting girls rejection of the physical science subjects. It was worded, "How could the physical sciences be made more attractive for girls?" This appendix outlines the range of replies that were received from the teachers.

Better careers advice was mentioned by many teachers to be an important step that would encourage more girls to continue with the physical sciences. There was a general feeling that girls are largely unaware of the importance of physics for a range of careers which could interest them.

> Perhaps a better appreciation of what jobs are available if they take the physical sciences.
>
> Female biology teacher (12)

Some teachers suggested tinkering with the science syllabi and introducing topics that would be of greater interest to girls. However, there was conern that such action would present girls with a false image of the subject.

> I think if we try to draw people into studying a particular subject, we must give them a realistic picture of what we are offering them in the future, and not just attempt to drag them in with some sort of glossy idea-catching course which bears no relationship to the subject thereafter. That doesn't mean we shouldn't do everything to try and make it interesting, but we must be honest in what we do in those directions.
>
> Male chemistry teacher (3)

Other teachers were concerned that the introduction of different topics into a syllabus would adversely alter the objectives and nature of the subject being taught.

> I've looked at courses which are called science for girls, which concentrate on dry cleaning and babies nappies and all this sort of nonsense. I must admit I've got no time for that. It won't actually work.
> (Interviewer: Why?)
> Because I think it's dragging the subject down to a very restricted

level, rather than bringing them towards a wider view of things. I won't entertain that sort of thing. I suppose we could introduce so called girls' interests in a small way. I don't really feel inclined to do it. I want to stick to a broad course. So really it's a matter of encouraging and cajoling the girls to realise that they are not restricted, there's no reason why they shouln't know about the workings of the car engine or where the electricity supply comes from. It's a slow matter of changing attitudes.

Male general science teacher (11)

The last quote introduces another remedy to the 'Girls and Science' problem. It is that the girls themselves must change. If only we could improve their attitudes towards the physical sciences, then they would be more interested and more eager to study the physical science subjects. Other respondents suggested that the teachers could be instrumental in instigating improved participation amongst girls.

Positive efforts by the teacher to make it clear to girls that they are expected to participate and be successful.

Male chemistry teacher (23)

Some teachers pointed to the ratio of male to female physical science teachers. There are too few female teachers to act as role models and offer encouragement to girls.

Have more women teachers.

Male physics teacher (8)

Instead of seeking changes within just pupils or teachers, one respondent suggested that there should be attitude changes within the school as a whole.

If schools reflected a different atmosphere or different attitudes then I think we might see more girls taking the physical sciences, although it would be a very long process.

Male biology teacher (7)

Another respondent thought that the attitudes held within society at large needed to be changed.

We've got to re-educate our society to stop thinking in terms of the boys being mechanics and the girls being nurses. Alternatively, the school can make general science, integrated science, physical science, whatever you want to call it, not optional but compulsory. And then once they have not got the choice, they will accept that as the course, in just the same way that they don't kick up at doing maths.

Male physics teacher (9)

The above quote also puts forward another solution - compulsory science for all, presumably up to the age of 16.

The last group of suggestions all refer to the way in which a subject is taught, in particular the way in which material i presented. Teachers thought that steps such as introducing single sex groups, making greater use of female-related examples, and decreasing the mathematical content of a course might well encourage more girls to study the physical science subjects.

> In some cases, probably being willing to take them actually in groups by themselves. I don't like single sex teaching, but I think perhaps in some cases it may well encourage some of the girls to take up physics or chemistry if they feel that they're not showing themselves up in front of the boys.
>
> Female biology teacher (12)

> If we keep maths to a minimum early on, and try and build in more essay type answers which girls tend to do better at. Also I must admit I tend to relate things we do in the lesson to the outside world, I do through my experience which is male experience. So if I'm talking about energy transfer I tend to talk about things like bullets rather than hair dryers or something like that.
>
> Male physics teacher (5)

In summary, the teachers suggested four main approaches to encourage more girls to study the physical science subjects.

(i) Provide girls with better careers advice.

(ii) Change the science syllabi.

(iii) Change the way that science is taught.

(iv) Change attitudes - of the girls, their teachers, the schools, society.

The first three suggestions are fairly predictable responses. These remedies are routinely presented in educational journals, magazines and newspapers (Duxbury, 1984; Harding, 1982; Hearn, 1979). Besides their accessibility in the educational press, these remedies, which in their implementation all entail obvious actions and changes, are relatively easy to conceive and accept. However, the last approach suggested in the list is more remarkable, for it entails subtle changes that are difficult to implement or even detect, but which nevertheless can have very far reaching effects. Kelly et al. (1984) report that even teachers who were engaged in a research project (the GIST project) to encourage girls into science and technology, were generally unaware of the important contribution that attitude change can make to attacking the root of the problem.

APPENDIX 6.1

SCALE CONSTRUCTION

Any measuring device should ideally possess the following characteristics: unidimensionability, linearity and equal intervals, reliability and validity (Oppenheim, 1966). The scale should be sufficiently sensitive to detect small differences in the property being measured. For rating scales, it is also important that irrelevant variables do not affect the measurement made, and that the measurement is equivalent to others made by alternative instruments measuring the same property (Smith, 1975). In addition, it is desirable that rating scales should be easy to construct, administer, score and interpret.

A scale is unidimensional if it measures a single, specific attribute. Personal attributes, e.g. perceptions, beliefs, attitudes, are frequently measured by a composite score or index. A composite score is made by combining, often by straightforward addition, the responses to several items to form a single index of the underlying construct being measured. If the index refers to a single dimension or aspect of the construct, then it is unidimensional. In this study, the Masculinity Indices and the Females' Social Roles scale are all unidimensional.

Various procedures are available to enable scales to be developed that possess desired characteristics, e.g. unidimensionability. Item analysis can be used to maximize internal consistency within the items, but this method will not detect the presence of more than one dimension in the item set. Cluster analysis provides the simplest technique for investigating the dimensionality of a set of items (see Appendix 6.6). A more sophisticated statistical method for producing a unidimensional scale is offered by factor analysis (see Appendix 6.7).

A linear scale is one that follows the straight-line model. If it is also an interval scale, then the intervals between the scale points are equal, but the scale has no true zero point. Attitude scales assume the straight-line model, and very often assume equal intervals as well. It is doubtful whether the equal interval assumption is absolutely accurate. As Oppenheim (1966) observes "... Numerically similar attitude-scale differences may represent very different psychological distinctions" (p.121). Strictly speaking attitude scales are most probably ordinal scales, like rating scales. An ordinal scale is one which allows the individuals in a data set to be ranked in order but which says nothing about the distance between the categories or ranks. The rating scales used in the STOSS questionnaire are examples of ordinal scales. Unfortunately the nonparametric statistical techniques available for use with ordinal scales are limited and of low power, so parametric statistics are frequently used (see section 4.6.1).

The validity of a subject's responses can be affected by the presence of ambiguous statements or items, ambiguous scales, scales which confound two or more different dimensions, or the tendency of the subject to operate response sets (Dawes, 1972). A response set is a systematic set or bias which pervades a subject's answers to a questionnaire and results in response bias, a situation where responses indicate something other than what the rating scale was intended to measure. A number of response sets have been identified, e.g. the tendency to be lenient, to give an average or central rating, to give similar ratings to items which seem logically related, to give similar ratings to adjacent items in the rating scale (proximity error), to be influenced by the general impression of the object being rated (halo effect), to be influenced by the wording of the item (semantic bias), and to give a socially desirable response (Guilford, 1954). Some of these response sets, especially the last two examples, together with the tendency to answer positively to questions (acquiescence response set), are applicable to attitude scales.

A number of approaches to control response biases have been suggested (Guilford, 1954; Smith, 1975). Most of them involve improving or modifying the questionnaire. During the construction of the attitude and rating scales used in this study due regard was paid to the problem of systematic response biases and to the range of methods available to counteract them. It is particularly important to control response biases because they affect the validity of the findings. Validity, along with reliability, are discussed in sections 6.3 and 6.4.

APPENDIX 6.2

QUESTIONNAIRE DESIGN

The design of questionnaires is guided by a number of criteria, including their relevance, their comprehensiveness, their comprehensibility, their feasibility, and their appearance (Open University, 1979). Relevance to the research problem is the basic criterion in deciding what questions or scales should be included in a questionnaire. It is important to ensure that the contents of a questionnaire adequately cover the various aspects of the research problem and will produce all the data required in a form suitable for the planned data analysis. If it is the intention to use parametric statistics, as it was in this study, then measurements should be made on scales that approximate to interval scales.

For a questionnaire to be comprehensible, it must be understood by and meaningful to all people who have to read it. A basic requirement of an effective questionnaire is that it should consist of questions which give "maximum opportunity for complete and accurate communication of ideas between the researcher (or interviewer) and the respondent" (Cannell & Kahn, 1968, p.553). They suggest that the communication process depends upon three components: language, frame of reference, and conceptual level of the questions. The language used should not be too complex or too condescending. The conceptual level of the questions should not be too difficult and the questions should, as far as is possible, be expressed within the respondents' frame of reference. It helps to try and see issues through the eyes of the respondents before phrasing questions. Much useful advice concerning the wording of questions and attitude statements has been offered over the years (Edwards, 1957; Moser & Kalton, 1971; Oppenheim, 1966; Parten, 1950; Payne, 1951). By adhering to the advice, it was hoped that many of the usual problems of ambiguity, bias, complexity, vagueness, emotional content, etc. could be largely avoided in this investigation.

Questionnaires should not make unreasonable demands upon respondents. They should not be too complex or take too long to fill. The feasibility of a questionnaire is largely determined by its content, but organisation and appearance can also contribute. Thought needs to be given to the order in which questions are asked (Oppenheim, 1966; Parten, 1950; Smith, 1975; Youngman, 1978). Easy questions are generally asked at the beginning to encourage commencement of the task, without taxing or threatening the respondent. Subsequent questions should be arranged in some form of logical groupings. But it should not be forgotten that replies to later questions might be influenced by earlier questions. Attention was paid to this problem in the sequencing of some of the questions and scales in the present study. Consideration was also given to the overall organisation of the questionnaires and their appearance. Besides encouraging subjects to respond favourably to the questionnaire and so increasing the response rate, a well laid out questionnaire will facilitate both the completion and analysis stages. Several writers have offered advice concerning the format and appearance of questionnaires and this was duly heeded (Henerson et al., 1978; Parten, 1950; Youngman, 1978).

APPENDIX 6.3

McKENNELL'S ATTITUDE SCALING PROCEDURE

McKennell (1970) proposed the use of coefficient alpha with cluster or factor analysis as a general strategy for attitude scale construction. The method facilitates the construction of scales that are both short and homogeneous, but also representative of the attitude domain and reliable. Although the technique involves a number of procedures, they are all relatively straightforward and simple. The various steps are briefly outlined below.

1. Identify the attitude domain.
2. Conduct exploratory research to clarify the content of the domain and to discover different aspects of the domain. Unstructured interviews and group discussions are commonly used research methods at this stage.
3. From the opinions expressed in the exploratory research, assemble a selection of items and construct a scale for used in a pilot study.
4. Present the items to a pilot sample.
5. Construct a correlation matrix from the data collected.
6. Use cluster analysis (see Appendix 6.6 for details of elementary linkage analysis)to establish a hypothesis about the dimensionality of the set of attitude items.
7. Use factor analysis to refine the dimensionality hypothesis. Factor analysis helps to identify items that should be discarded from the measurement scale in order to make the underlying dimension(s) of the scale purer.
8. To obtain an estimate of the reliability of a scale formed from the items composing a cluster or factor calculate coefficient alpha (α). McKennell gives the following approximate formula for computing coefficient alpha.

$$\alpha = \frac{n\bar{r}_{ij}}{1 + (n-1)\bar{r}_{ij}}$$

where n = the number of items in the test

$\bar{r}_{ij}$ = the average of all the inter-item correlations

The above formula has the advantage of working directly from the correlations between the items.

The following formula is the one originally given by Cronbach (1951) to calculate alpha.

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\text{sum of variances of question scores}}{\text{variance of total test scores}}\right)$$

where n = the number of items in the test.

9. The effect upon the reliability of a scale of shortening it by removing a number of items can easily be calculated. Decisions regarding the length of a scale are usually guided by reliability considerations. Minimum alpha values of 0.60 are commonly adopted in survey research (Barker Lunn, 1969; Brynner, 1972).

When assessing the consequences for reliability of selecting a subset of items to represent a cluster, it is not necessary to solve the formula for each case. McKennell provides a table of values for alpha corresponding to different combinations of n and $\bar{r}_{ij}$.

10. Alpha can also be used to guide the selection of items when shortening a scale. For a given number of items, reliability is greatest when $\bar{r}_{ij}$ is at a maximum. $\bar{r}_{ij}$*, the average correlation of each item with all the other items in the cluster, is calculated for each item since this value indicates each item's contribution to $\bar{r}_{ij}$. The items are then ordered according to their $\bar{r}_{ij}$* values, and $\bar{r}_{ij}$ and alpha are also recorded. By discarding items with low $\bar{r}_{ij}$* values, the remaining items provide maximum possible $\bar{r}_{ij}$ and alpha values. Decisions concerning the number of items to be discarded are made with regard to $\bar{r}_{ij}$ and alpha values.

11. Selection of items should not be guided solely by adherence to rules. Validity considerations suggest that it would be unwise to select items that are very similar in content. It is better to produce a scale that shows a broader coverage of the attitude domain.

Table A6.4/1 Additional construct validation techniques and their use

(Numbers refer to section of the thesis where more details can be found)

Scale	Validation technique: Correlation	Validation technique: Cluster & factor analysis
1. School Subject Characteristics	6.6.1.4	
2. Masculinity Index	6.6.2.2.2	6.6.2.2.1
3. Characteristics of Science		7.1.3.2
4. Scientist Stereotypes	8.3.2.1	
5. Preference for Subject Characteristics	8.3.2.3	
6. Females' Social Roles		6.7.3.2
7. Reasons for Success/Failure		7.3.1.3 & 7.3.2.3
8. Reasons for Choosing/Dropping		7.3.3.4 & 7.3.4.4
9. Marking Exercise		6.9.1.6

APPENDIX 6.4

ADDITIONAL CONSTRUCT VALIDATION TECHNIQUES EMPLOYED

Henerson et al. (1978) and Open University (1979) list a number of techniques for defending the construct validity of a measurement.

1. Opinions of judges. Experts are shown the scale and asked to comment upon its purpose. Their conclusions should agree, and coincide with the construct that the scale is designed to measure.

2. Criterion group studies. People judged independently to possess the construct being measured should score higher than people not possessing the construct.

3. Correlations. Respondent' scores from the scale should correlate with their scores from another measure of the same or a related construct.

4. Correlation matrices, e.g. Campbell and Fiske's multitrait-multi-method matrix (1959) which investigates convergent and discriminant validity at the same time by the use of a matrix of correlation coefficients between different measures.

5. Cluster and factor analysis. These techniques allow the dimensionality of a scale to be assessed. Furthermore factor analysis, besides indicating those items which measure the same construct, also indicates to what extent they measure that construct.

6. Internal consistency of the scale. Discussed in section 6.4.

7. Test-retest stability. Discussed in section 6.4.

8. Fair administration. Discussed in section 6.4.

9. Studies of responses to the scale. Discussed in section 6.4.

The last four construct validation techniques in the above list were used extensively in the present study and have been reported fully in section 6.4. Several of the other techniques were also used, but they were not universally applied to all the scales. Table A6.4/1 shows the use made of these additional validation techniques.

APPENDIX 6.5

RESULTS OF THE PRELIMINARY STUDIES TO DETERMINE ADJECTIVE PAIRS' GENDER CONNOTATIONS

Table A6.5/1 Study 1, Sample B (N=18)

	M	F		M	F
Hard	18	0	Soft	0	18
Weak	2	16	Powerful	16	2
Tender	1	17	Tough	17	1
Cold	14	4	Warm	4	14
Intimate	3	15	Remote	15	3
Light	0	18	Heavy	18	0
High	11	5	Low	5	11

Table A6.5/2 Study 2, Sample D (N=49)

	M	F		M	F	Refused
Active	20	6	Passive	6	20	23
Hard	26	7	Soft	7	26	16
Weak	4	20	Powerful	20	4	25
Tender	1	25	Tough	25	1	23
Cold	16	4	Warm	4	16	29
Intimate	1	16	Remote	16	1	32
Light	2	19	Heavy	19	2	28
High	9	8	Low	8	9	32

Table A6.5/3 Study 3, Sample E (N=24)

	Extremely	Quite masculine	Slightly masculine	Neutral	Slightly feminine	Quite feminine	Extremely feminine
Tender				9	9	4	2
Remote		1	6	16			1
Passive				17	3	4	
Soft				8	8	5	3
Cold		2	5	17			
Weak			2	16	5	1	
Low				24			
Heavy	1	5	2	15	1		
Hard	1	7	2	13	1		
Warm				15	7	2	
High	1		1	20	1		1
Powerful	5	6	5	8			
Intimate				15	6	2	1
Light				16	7	1	
Active		3	4	17			
Tough	4	9	5	6			

Table A6.5/4 Study 4, Sample J (N=22)

	Very masculine (3)	(2)	(1)	(-1)	(-2)	Very feminine (-3)	Refused
Tender				11	6	5	
Remote	2	1	9	8			2
Passive			2	12	3	2	3
Soft			1	12	6	3	
Cold	1	1	10	5	2		3
Weak	1	2	2	12	2	2	1
Heavy	3	6	10	2	1		
Hard	2	5	11	1		1	2
Warm			3	9	2	4	4
Powerful	8	6	6		1		1
Intimate		1		11	5	2	3
Light				10	9		3
Active	4	4	10	1	2		1
Tough	11	3	5				3

Allocating the values to the ratings as shown at the head of the columns (refused = 0), mean ratings were calculated for each adjective.

Table A6.5/5 Mean ratings allocated to each adjective

Active	1.14	Passive	-1.00
Hard	1.05	Soft	-1.45
Powerful	1.82	Weak	-0.59
Tough	2.00	Tender	-1.73
Cold	0.27	Warm	-1.00
Remote	0.41	Intimate	-1.14
Heavy	1.23	Light	-1.27

APPENDIX 6.6

ELEMENTARY LINKAGE ANALYSIS

McQuitty's (1957) elementary linkage analysis is a simple, rapid procedure for locating clusters of items in a correlation matrix. Even very large numbers of items can be allocated to clusters in only a few minutes. This speed is achieved at the expense of the method only being approximate. Each item is assigned to the cluster with which it has the highest single correlation. This will usually be the cluster with which it has the highest average correlation, but discrepancies can arise.

The procedure for assigning items to the clusters with which they have the highest single correlation involves the following steps.

1. Produce a correlation matrix for the whole set of items.
2. Mark the highest correlation in each column of the correlation matrix.
3. Extract the highest pair of correlations in the matrix. The two items involved constitute a reciprocal pair and will form the core of a cluster.
4. For each member of the reciprocal pair, search its row in the correlation matrix for other marked values. A marked value denotes that it is the highest correlation appearing in the column of a particular item. Add all such items to the cluster.
5. For each item added to the cluster in step 4, search its row in the correlation matrix for any other correlations that are marked as highest. Add any items thus identified to the cluster.
6. Continue to repeat step 5 for the most recently introduced items until no more marked correlations can be found. Then the cluster is complete.
7. Look for the highest pair of correlations among the remaining items, and repeat steps 4 to 6.
8. Repeat step 7 until all the items have been allocated to clusters.

APPENDIX 6.7

FACTOR ANALYSIS

Factor analysis is a method for determining the number and nature of the dimensions or factors that underlie a set of variables. Also, it indicates the strength of the relationship between each variable and each factor. Factor analysis is particularly useful in scale construction.

Factor analysis solutions recorded in the present study were obtained by application of the factor analysis programme available in the SPSS computer package (Nie et al., 1975). Following the advice of Kim & Mueller (1978), the programme's default option was used. This method provides principal factoring with iteration. The number of factors is determined by the number of the roots (eigenvalues) of the correlation matrix which are greater than or equal to 1.0. The factors are subjected to orthogonal rotation using the Varimax method.

Only limited use was made of factor analysis, because few of the data sets obtained in the study were sufficiently large to produce reliable results. Large sample size is required to minimise the standard error of correlations and to reliably identify factors and factor loadings. Small samples can produce error factors and distort the true factor structure. Cattell (1952) suggested that a ratio of respondents to variables of 4:1 may be adequate, but Kerlinger (1973) recommends a ratio of 10:1. Comrey (1978) and Loo (1983) suggest that an acceptable sample should consist of at least 200 respondents. Since the major concern is to obtain stable correlations, the ratio of respondents to variables is largely unimportant (Loo, 1983).

In this study, factor analysis has been used with sample sizes smaller than those recommended above, e.g. in the development of the Females' Social Roles attitude scale. However, each time the results obtained from factor analysis were checked and confirmed using alternative methods, e.g. cluster analysis, item analysis. Where factor analysis has been used as the only method of analysis, e.g. to determine the factors underlying teachers' marking practices, the sample size was always greater than 200.

Table A6.8/1 Characteristics possessed by physical and biological science

Characteristic	Physical science (N=34)	Biological science (N=23)
Practical	1.44	1.80
Technical	1.71	2.52
Applied	1.87	2.13
Unfamiliar	2.31	3.15
Dirty	3.00	3.24
Mechanical	2.71	3.44
Inanimate	2.15	3.74
Complex	2.00	1.96
Conceptual	1.38	1.94
Analytical	1.52	2.28
Objective	1.41	2.04
Experimental	1.41	1.70
Pupil-centred	2.18	2.28
Teacher-centred	2.37	2.46
Structured	1.56	1.72
Syllabus-bound	1.56	1.70
Factual	1.62	1.52
Routine	2.77	2.65
Descriptive	2 38	2.09
Wordy	2.79	1.94
Feminine	3.19	3.30
Speculative	2.50	2.65
Exploratory	1.82	1.87
Open-ended	2.65	2.50
Syllabus-free	3.67	3.44

APPENDIX 6.8

CHARACTERISTICS POSSESSED BY PHYSICAL AND BIOLOGICAL SCIENCE

The table below shows mean ratings of the characteristics possessed by physical science and biological science subjects. A rating of 1 denotes that the subject is very practical, technical, etc., whilst a rating of 4 indicates that the subject is not at all practical, technical, etc. Thus a low mean rating implies that science teachers regard the subject as possessing that characteristic.

Characteristic	Physical science	Biological science
Observational	1.32	1.79
Creative	2.68	2.78
Concerned with people	2.97	1.96
Concerned with objects	1.82	3.09
Concerned with social issues	2.95	2.36
Precise	1.80	2.11
Mathematical	1.91	2.65
3-dimensional	2.02	2.13
Difficult	1.93	1.78
Dangerous	2.93	3.11
Abstract	1.95	2.71
Academic	1.79	2.00
Relevant for careers	1.35	2.13
Relevant for family life	2.62	1.61
Relevant for everyday life	2.18	1.78
Theoretical	1.82	2.17
Logical	1.38	2.02
Masculine	2.72	3.30
Convergent	2.43	3.28
Impersonal	2.32	3.11
Concise	2.02	2.44
Demanding	1.56	1.83
Interesting	1.60	1.11
Animate	3.09	1.52

APPENDIX 6.9

THE OPINIONS QUESTIONNAIRE

IMAGE OF SCIENCE SUBJECTS

Please complete the 'Characteristics of School Subjects' before answering this questionnaire.

PLEASE SUPPLY
(For statistical purposes only)

Your name . ____________________

Sex: Male ☐
Female ☐

Age: Under 30 ☐
30 - 39 ☐
40 - 49 ☐
50 and over ☐

Your principal teaching subject ____________________

Subsidiary teaching subject(s) ____________________

This questionnaire refers to science subjects as they are taught in secondary schools up to CSE/O level standard.
Answer by placing a tick in the appropriate box.

1. Do you think that the general public regards the physical science subjects (e.g.Physics and Chemistry) to be
   - masculine subjects ☐
   - feminine subjects ☐
   - neutral subjects ☐

2. Do you think that the general public regards the biological science subjects to be
   - masculine subjects ☐
   - feminine subjects ☐
   - neutral subjects ☐

3. Do you yourself generally regard the physical science subjects to be
   - masculine subjects ☐
   - feminine subjects ☐
   - neutral subjects ☐

4. Do you yourself generally regard the biological science subjects to be
   - masculine subjects ☐
   - feminine subjects ☐
   - neutral subjects ☐

SECTION A Public Opinion

According to public opinion, do any of the following factors give the physical science subjects a masculine image?
Answer by circling the appropriate response. Only circle '?' if absolutely uncertain. Please answer this section before proceeding to the next section.

1. Scientific language	Yes	?	No
2. Vocabulary	Yes	?	No
3. Content	Yes	?	No
4. Examples used	Yes	?	No
5. Analogies used	Yes	?	No
6. School textbooks	Yes	?	No
7. Story books	Yes	?	No
8. Comics	Yes	?	No
9. Films	Yes	?	No
10. Television	Yes	?	No
11. Adverts	Yes	?	No
12. Radio	Yes	?	No
13. Social pressure	Yes	?	No
14. Man's role in society	Yes	?	No
15. Instruments	Yes	?	No
16. Apparatus	Yes	?	No
17. Equipment	Yes	?	No
18. Techniques	Yes	?	No
19. Manual skills	Yes	?	No
20. Intellectual abilities required	Yes	?	No
21. Personality characteristics required	Yes	?	No
22. Number of male scientists	Yes	?	No
23. Stereotyping	Yes	?	No
24. Impersonality of science	Yes	?	No
25. Lack of social concern	Yes	?	No
26. Mathematical component	Yes	?	No
27. Mechanical aspects	Yes	?	No
28. Practical subjects	Yes	?	No
29. Dangers	Yes	?	No
30. 'Dirty hands'	Yes	?	No
31. Complex subject	Yes	?	No
32. Analytical subject	Yes	?	No
33. Historical factors	Yes	?	No
34. Number of male pupils	Yes	?	No
35. Teachers' attitudes	Yes	?	No
36. Number of male teachers	Yes	?	No
37. Teaching style	Yes	?	No
38. Science syllabuses	Yes	?	No
39. Single sex schools	Yes	?	No
40. Tradition	Yes	?	No
41. Employment prospects	Yes	?	No
42. Prestige of science	Yes	?	No
43. Power of science	Yes	?	No

SECTION B Personal Opinion

In your opinion do any of the following factors give the physical science subjects a masculine image?
Again answer by circling the appropriate response. Only circle '?' if absolutely uncertain.

1.	Scientific language	Yes	?	No
2.	Vocabulary	Yes	?	No
3.	Content	Yes	?	No
4.	Examples used	Yes	?	No
5.	Analogies used	Yes	?	No
6.	School textbooks	Yes	?	No
7.	Story books	Yes	?	No
8.	Comics	Yes	?	No
9.	Films	Yes	?	No
10.	Television	Yes	?	No
11.	Adverts	Yes	?	No
12.	Radio	Yes	?	No
13.	Social pressure	Yes	?	No
14.	Man's role in society	Yes	?	No
15.	Instruments	Yes	?	No
16.	Apparatus	Yes	?	No
17.	Equipment	Yes	?	No
18.	Techniques	Yes	?	No
19.	Manual skills	Yes	?	No
20.	Intellectual abilities required	Yes	?	No
21.	Personality characteristics required ...	Yes	?	No
22.	Number of male scientists	Yes	?	No
23.	Stereotyping	Yes	?	No
24.	Impersonality of science	Yes	?	No
25.	Lack of social concern	Yes	?	No
26.	Mathematical component	Yes	?	No
27.	Mechanical aspects	Yes	?	No
28.	Practical subjects	Yes	?	No
29.	Dangers	Yes	?	No
30.	'Dirty hands'	Yes	?	No
31.	Complex subject	Yes	?	No
32.	Analytical subject	Yes	?	No
33.	Historical factors	Yes	?	No
34.	Number of male pupils	Yes	?	No
35.	Teachers' attitudes	Yes	?	No
36.	Number of male teachers	Yes	?	No
37.	Teaching style	Yes	?	No
38.	Science syllabuses	Yes	?	No
39.	Single sex schools	Yes	?	No
40.	Tradition	Yes	?	No
41.	Employment prospects	Yes	?	No
42.	Prestige of science	Yes	?	No
43.	Power of science	Yes	?	No

COMMENTS

Do feel free to add any further factors which you believe to be important but which have not been listed, or to write comments about the idea of assigning gender to school subjects.

Margaret Spear

APPENDIX 6.10

QUESTIONS ASKED ABOUT THE WRITTEN WORK OF BOYS AND GIRLS

WRITTEN WORK OF BOYS AND GIRLS

I would also like to inquire about the characteristics of the written work of secondary school pupils. In particular, I am interested in any features which distinguish the written work of a girl from that of a boy.

1. Would you say that you can generally distinguish between the written work of boys and girls?

   Yes ☐ No ☐

2. Can you briefly indicate any features that you consider to be typical of the written work of girls and boys.

Features typical of girls' written work	Features typical of boys' written work

3. What is your main teaching subject? ______________
4. Please indicate
   - (a) Your sex:
     - Male ☐
     - Female ☐
   - (b) Your teaching experience:
     - Less than 2 years ☐
     - 2 - 5 years ☐
     - 5 - 10 years ☐
     - 10 - 20 years ☐
     - Over 20 years ☐

Thank you for your cooperation and help.

Margaret Spear

APPENDIX 6.11

DETAILED PILOT RESULTS FROM THE FEMALES' SOCIAL ROLES QUESTIONNAIRES

Table A6.11/1 Female Role scale items and science teachers' responses (percentage frequency) (First pilot, N=45)

	Strongly Disagree	Mildly Disagree	Mildly Agree	Strongly Agree
1. A good mother would not go out to work whilst she had a child under 5.	4.4	20.0	33.3	42.2
2. The only really satisfying role for a woman is as a wife and mother.	73.3	24.4	2.2	0
3. Looking after children is just as much the father's job as the mother's.	0	8.9	28.9	62.2
4. Women are as good as men at complicated technical matters.	2.2	46.7	24.4	26.7
5. Girls should be encouraged to be ambitious in terms of a career.	2.2	0	20.0	77.8
6. A man should not be expected to look after a baby under normal circumstances.	46.7	35.6	13.3	4.4
7. Women are not suited to jobs of great stress and responsibility	62.2	22.2	13.3	2.2
8. Women are men's equals intellectually.	8.9	4.4	15.6	71.1
9. A man should be responsible for providing money for his wife's personal use even if she is capable of earning it herself.	24.4	31.1	33.3	11.1
10. Women's most important job is to look after the comforts of men and children.	57.8	24.4	13.3	4.4
11. A woman should allow her husband to feel superior even if this involves belittling herself.	77.8	6.7	13.3	2.2
12. Women should be happy to take second place to their husbands.	68.9	17.8	11.1	2.2
13. Women should obey their husbands.	66.7	13.3	15.6	4.4

	Strongly Disagree	Mildly Disagree	Mildly Agree	Strongly Agree
14. A situation in which a women works whilst a man stays at home and looks after the children is not right	42.2	22.2	31.1	4.4
15. A woman should be quite willing to give up her own job if her husband can gain promotion by moving to another area.	6.7	35.6	44.4	13.3
16. A woman's career is not as important as a man's.	42.2	15.6	35.6	6.7
17. Femininity is a woman's greatest attribute.	17.8	20.0	35.6	26.7
18. The age at which women qualify for a retirement pension should be the same as for a man.	2.2	0	15.6	82.2
19. It is the man's job to make the major decisions.	42.2	28.9	15.6	13.3
20. A woman could not reach the top in her career without her family suffering.	15.6	31.1	28.9	24.4
21. If a child is ill then it is the mother's duty rather than the father's to take time off work to look after him/her.	13.3	33.3	33.3	20.0
22. Women should only have children if they are prepared to give up their jobs to look after them until they are old enough to go to school.	6.7	22.2	26.7	44.4
23. Women should feel uninbibited about taking the sexual initiative.	2.2	2.2	26.7	68.9
24. The saying "a women's place is in the home" is generally correct.	44.4	26.7	26.7	2.2
25. A woman can be a good wife and mother even if she has a very demanding job.	6.7	24.4	37.8	31.1

Table A6.11/2 FEM scale items and science teachers' responses (percentage frequency) (Second pilot, N=49)

	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
1. Women have the right to compete with men in every sphere of activity.	55.1	30.4	2.0	8.2	4.1
2. As head of the household, the father should have final authority over his children.	8.2	18.4	10.2	16.3	46.9
3. The unmarried mother is morally a greater failure than the unmarried father.	0	2.0	14.3	18.4	65.3
4. A woman who refuses to give up her job to move with her husband would be to blame if the marriage broke up.	6.1	16.3	28.6	28.6	20.4
5. A woman who refuses to bear children has failed in her duty to her husband.	6.1	16.3	12.2	22.4	42.9
6. Women should not be permitted to hold political offices that involve great responsibility.	0	6.1	2.0	18.4	73.5
7. A woman should be expected to change her name when she marries.	6.1	22.4	28.6	14.3	28.6
8. Whether or not they realise it, most women are exploited by men.	14.3	14.3	28.6	30.6	12.2
9. Women who join the Women's Movement are typically frustrated and unattractive people who feel they lose out by the current rules of society	2.0	18.4	26.5	32.7	20.4
10. A working woman who sends her six month old baby to a day nursery is a bad mother.	16.3	16.3	14.3	24.5	28.6

	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
11. A woman to be truly feminine should gracefully accept chivalrous attentions from men.	2.0	20.4	30.6	26.5	20.4
12. It is absurd to regard obedience as a wifely virtue.	44.9	28.6	10.2	8.2	8.2
13. The "clinging vine" wife is justified provided she clings sweetly enough to please her husband.	4.1	14.3	42.9	16.3	22.4
14. Realistically speaking, most progress so far has been made by men and we can expect it to continue that way.	2.0	42.9	16.3	22.4	16.3
15. One should never trust a women's account of another woman.	2.0	10.2	20.4	38.8	28.6
16. It is desirable that women who join the police force should undertake exactly the same duties as men.	8.2	20.4	24.5	38.8	8.2
17. Women are basically more unpredictable than men.	6.1	24.5	22.4	26.5	20.4
18. It is all right for women to work but men will always be the basic breadwinners.	6.1	30.6	10.2	30.6	22.4
19. A woman should not expect to go to the same places or have the same freedom of action as a man.	0	0	6.1	36.7	57.1
20. Swearing generally sounds worse coming from a woman.	10.2	53.1	14.3	8.2	14.3

Table A6.11/3 Selection criteria for Female Role scale items, first pilot

Item No.	Mean	S.D.	Factor loading	Item-whole r	$(\bar{H} - \bar{L})$
1	1.11	1.28	0.443	.47	1.58
2	3.69	0.60	0.624	.63	0.92
3	3.44	0.89	0.468	.46	1.08
4	2.27	1.36	0.318	.39	1.33
5	3.71	0.70	-0.144	-.10	-0.25
6	3.07	1.20	0.396	.42	0.92
7	3.29	1.14	0.331	.38	1.08
8	3.36	1.26	0.440	.48	1.42
9	2.24	1.43	0.646	.63	2.17
10	3.18	1.23	0.766	.73	2.08
11	3.44	1.16	0.744	.72	1.75
12	3.40	1.10	0.761	.74	1.83
13	3.22	1.30	0.676	.68	2.17
14	2.67	1.41	0.554	.58	2.08
15	1.78	1.26	0.661	.68	2.08
16	2.51	1.50	0.732	.74	2.75
17	1.67	1.51	0.406	.41	1.58
18	3.76	0.68	-0.188	-.15	-0.17
19	2.71	1.49	0.648	.64	2.17
20	1.84	1.49	0.464	.53	2.33
21	1.87	1.42	0.653	.69	2.67
22	1.20	1.39	0.488	.53	1.92
23	3.58	0.81	0.013	.05	0
24	2.84	1.31	0.724	.73	2.58
25	2.62	1.34	0.580	.61	2.33

$(\bar{H} - \bar{L})$ Mean of high scorers - mean of low scorers

Table A6.11/4 Selection criteria for FEM scale items, second pilot

Item no.	Mean	S.D.	Factor loading	Item-whole r	$(\bar{H} - \bar{L})$
1	4.25	1.11	0.396	.43	1.00
2	3.76	1.42	0.719	.74	2.67
3	4.47	0.82	0.513	.52	1.17
4	3.41	1.17	0.702	.71	2.25
5	3.80	1.32	0.735	.74	2.75
6	4.59	0.81	0.447	.47	0.92
7	3.37	1.29	0.744	.74	2.25
8	2.88	1.24	0.421	.49	1.75
9	3.51	1.08	0.445	.50	1.33
10	3.33	1.46	0.563	.61	2.17
11	3.43	1.10	0.411	.43	1.33
12	3.94	1.28	0.363	.43	1.25
13	3.39	1.12	0.384	.40	1.00
14	3.08	1.19	0.522	.53	1.83
15	3.82	1.03	0.526	.53	1.50
16	2.82	1.11	0.048	.15	0.33
17	3.31	1.23	0.512	.52	1.42
18	3.33	1.30	0.676	.65	2.16
19	4.51	0.62	0.511	.50	0.75
20	2.63	1.22	0.492	.49	1.42

$(\bar{H} - \bar{L})$ Mean of high scorers - mean of low scorers

APPENDIX 6.12

RESULTS OF THE PRELIMINARY STUDIES TO CHOOSE NAMES

Table A6.12/1 Study 1, Sample B (N=20)

Name	Popularity: Like	Uncertain	Dislike
Sally Smith	13	1	6
Simon Smith	6	3	11
Susan Clark	11	3	6
Dave Clark	7	1	12
Mary Williams	8	6	6
Paul Williams	14	3	3
Jane Brown	9	4	7
John Brown	8	2	10
Ann Jones	14	2	4
Alan Jones	14	2	4
Pam Davis	10	2	8
Peter Davis	12	4	4
Cathy Taylor	6	5	9
Colin Taylor	10	4	6

Table A6.12/2 Study 2, Sample D (N=49)

Name	Popularity		
	Like	Uncertain	Dislike
Susan	26	6	17
Linda	22	6	21
Claire	33	5	11
Sarah	35	4	10
David	42	2	5
John	30	8	11
Stephen	29	8	12
Mark	36	3	10
Paul Williams	26	11	12
Sarah Williams	26	8	15
Alan Jones	23	12	14
Anne Jones	20	9	20
Mark Williams	24	12	13
Joanne Williams	22	15	12
Alison Jones	19	18	12
Christopher Jones	24	12	13

Table A6.12/3 Study 3, Sample E (N=24)

	Popularity		
Name	Like	Uncertain	Dislike
Susan	12	4	8
Claire	15	3	6
Sarah	19	4	1
Elizabeth	14	3	7
Helen	15	5	4
Jane	15	3	6
Margaret	8	8	8
Mary	5	6	13
David	17	5	2
John	10	8	6
Peter	11	5	8
Michael	16	6	2
Stephen	11	6	7
Mark	12	4	8
Paul	13	4	7
Andrew	11	8	5

Table A6.12/4 Study 4, Sample J (N=34)

Name	Popularity		
	Like	Uncertain	Dislike
Claire	24	5	5
Sarah	23	4	7
Nicola	19	7	8
Emma	25	5	4
Helen	21	4	9
Katherine	23	5	6
Emily	14	6	14
Charlotte	15	4	15
Rebecca	23	6	5
Elizabeth	16	9	9
Louise	19	3	12
Jane	19	8	7
Stephen	23	4	7
Mark	22	8	4
Paul	25	2	7
Andrew	22	6	6
David	23	2	9
Richard	22	5	7
Matthew	25	4	5
Daniel	19	6	9
Christopher	22	5	7
Michael	19	6	9
Alex	15	3	16
John	19	7	8

Table A6.13/1 Ratings and rankings awarded to unaccompanied (N=45) and accompanied (N=339) homework essays

	Essay 1			Essay 2			Essay 3		
	Rank	$\bar{x}$	s.d.	Rank	$\bar{x}$	s.d.	Rank	$\bar{x}$	s.d.
Merit									
Alone (standard)	2			1			3		
With write-up (O level suitability)	2			3			1		
Attitude									
Alone	1	3.84	0.80	2	3.53	1.10	3	3.07	1.03
With write-up	2	3.64	0.76	3	3.51	0.80	1	3.75	0.80
Interest									
Alone	1	3.89	0.75	2	3.27	1.18	3	2.98	0.97
With write-up	2	3.69	0.80	3	3.24	0.78	1	3.71	0.80

APPENDIX 6.13

HALO EFFECTS IN THE MARKING EXERCISE

During a preliminary study (sample K) the homework essays were rated on their own, without accompanying experimental write-ups. The same homework essays, together with experimental write-ups, constituted the marking exercise in the final data collection stage (sample P). The ratings given to the essays under these two different administration conditions, have been compared to see whether halo effects were operating. Halo is generally defined as the tendency to allow an estimation of one characteristic of a person to influence the estimation or rating of another characteristic of that person.

The original objective had been to select three essays of average standard. However, the preliminary study revealed that although a majority of the subjects judged two of the essays to be of average standard, one was considered to be above average standard (see section 6.9.1.5). During the marking exercise, in which each essay followed an experimental write-up, the respondents were not asked specifically to judge the standard of the essays, and so no direct comparison between the two conditions is possible. However, respondents were asked to judge each pupil's suitability for O level physical science courses. Presumably this judgement was made on the basis of the standard reached by the child on both pieces of work. Table A6.13/1 shows that the ranking of the unaccompanied essays on standard is not identical to the ranking of the write-ups plus essays on O level suitability. But, the ranking of the sample pairs on O level suitability was identical to the ranking of the write-ups on standard. Although standard and O level suitability are not equivalent rating variables, and although the evidence on which judgements were made was not identical in the two conditions, the results still point to the probability that the scientific content of the write-up contributed much more to the rating of O level suitability than did the general writing ability displayed in

Table A6.13/2 Ranking of homework essays for 'attitude towards science'

	Pupil 1	Pupil 2	Pupil 3
Standard of experimental write-up	Average	Below average	Above average
Standard of essay	Average	Above average	Average
Ranking of essay for 'attitude towards science'			
(i) Essay appearing without experimental write-up	1st	2nd	3rd
(ii) Essay appearing with experimental write-up	2nd	3rd	1st

the essay or the attitudes about science expressed in the essay.

The probable predominant influence of the standard of the write-up upon O level suitability ratings is neither unexpected nor unreasonable. However, the implications arising from the two direct comparisons which can be made are surprising. It appears that in the marking exercise, the teachers were influenced more by the standard of the preceeding write-up than by the contents of the essay when they judged a pupil's attitude towards science and interest in science, even though they had been directed to base their judgement on the essay alone. These relationships indicate that cognitive factors are more salient to judgements than affective factors, even when affective qualities are being judged. Table A6.13/2 illustrates the predominant influence of cognitive factors over affective factors in the assessment of attitude.

The above findings indicate that the halo effect is a significant variable in determining the assessment made of a piece of written work. If the first piece of work from a pupil is of a high standard, then the teacher is likely to judge subsequent work favourably too, and vica versa. This observation has important implications for teachers who attempt to ease their marking burden by marking two or more pieces of pupils' work at one time. If answers are not marked one at a time across the whole set, then there is a real danger that all the marks awarded to each pupil will be unduly influenced by the standard of the first answer or piece of work appearing in each pupil's book. Similar situations can arise when end of term examination papers are being marked. The simplest solution is to mark each question across the whole set of papers, before proceeding to the next question.

Finally, the results presented in Table A6.13/1 support the view that the standards of the three homework essays were not unacceptably dissimilar. Ratings of both of both attitude and interest for all three essays were overridden by signals conveyed by the preceeding experimental write-ups. This suggests that none of the essays created much impact upon the teachers, i.e. they were regarded as being ordinary or average, as was intended.

APPENDIX 6.14

INSTRUCTIONS FOR THE 'CARDS' EXERCISE

Part 2

Imagine that a week after taking up a new teaching post the Head of Science informs you that there is to be a Parents' Meeting the following week to discuss subject options for the fourth year. The parents expect to receive advice concerning their children's suitability for the different options. Consequently, the Head of Science asks you to provide him with a list of your third year pupils whom you would recommend should take the 'O' level course in your subject, and a list of pupils whom you judge to be below this examination standard.

When you consult your predecessor's record book you discover that his records for these pupils' progress in science are very fragmentary. However, to help out your Head of Science you agree to supply the lists, but only on the understanding that after you have become acquainted with the class you will be given an opportunity to review your assessment.

Please remove the cards from the large envelope. They represent your predecessor's records. After looking through the cards and assessing the information as far as you can, divide them into two piles - one pile for potential 'O' level candidates and another for pupils who are not suitable for this examination. Next indicate how confident you are of your decisions by writing

- '1' on the cards of those pupils that you are very confident are placed in the correct pile,
- '2' on those cards that you are fairly confident are placed in the correct pile, and
- '3' on the cards of those pupils whose allocation is somewhat uncertain.

Then put each pile into the appropriate small envelope, replace these small envelopes in the large envelope and write your name on it.

(I appreciate that this task may appear to be rather artificial but I am nevertheless interested in the decisions you make on the basis on the somewhat scant information available.)

Table A6.15/1 Allocation of each card and its median assessment (N=35)

	Pupil sex	Recommended for O level			Not recommended for O level			Median
		Very certain (1)	Quite certain (2)	Uncertain (3)	Uncertain (4)	Quite certain (5)	Very certain (6)	
1. Likes science	F	0	7	6	9	12	1	4.00
Must work harder	M	2	4	10	10	7	2	3.65
2. Average intelligence	F	4	8	2	7	8	6	4.00
Likes science	M	4	7	2	9	8	5	4.00
3. Likes science	F	0	0	0	5	13	17	5.46
Low marks	M	0	1	0	1	16	17	5.47
4. Tries hard	F	0	2	5	5	7	16	5.29
Dislikes science	M	0	0	3	7	13	12	5.08
5. Very intelligent	F	7	15	7	1	3	2	2.20
Dislikes science	M	9	12	8	2	3	1	2.21
6. High marks	F	5	13	7	4	6	0	2.46
Dislikes science	M	7	14	7	3	3	1	2.25
7. Tries hard	F	2	5	4	7	8	9	4.43
Average intelligence	M	3	7	3	6	7	9	4.25
8. Tried hard	F	0	0	0	3	7	25	5.80
Low marks	M	0	0	0	3	11	21	5.67
9. Very intelligent	F	14	18	2	0	0	1	1.69
Must work harder	M	17	18	0	0	0	0	1.53
10. High marks	F	11	18	4	1	1	0	1.86
Must work harder	M	9	21	4	0	1	0	1.90
11. Very intelligent	F	4	15	9	2	3	2	2.40
Low marks	M	3	13	10	4	2	3	2.65
12. Average intelligence	F	8	8	5	6	1	7	2.80
High marks	M	4	15	4	6	2	4	2.40

APPENDIX 6.15

EFFECT OF INFORMATION AVAILABILITY UPON TEACHER EXPECTATION

The validity of the Cards exercise was suspect on account of the cards having been sent to respondents through the post (see section 6.9.2.2). However, indications were that the validity had not, in fact, been seriously threatened, and for this reason the results are reported in this appendix. Still, their tentative nature should not be forgotten.

6.15.1.A Results

Analysis was based upon the responses of 35 teachers to 12 pairs of matched cards. This gave a total of 420 comparisons. Besides the rather small sample size, analysis was further restricted by the weak research design which did not include complete crossing of all the variables. A nonparametric method of analysis had to be used because the data had been collected using a disjunct ordinal measuring scale. Therefore, the criterion required for parametric tests of a continuous interval scale had not been met.

The influence of pupil sex upon individual teacher's expectations for the 12 card pairs, and upon the responses of the whole sample to individual card pairs was investigated by the Wilcoxon matched-pairs signed-ranks test (Ferguson, 1976). In no case was there justification for rejecting the null hypothesis. Thus individual teachers did not display significant sex bias, and the responses of the whole group of teachers to individual card pairs were not significantly sex biased. This overall lack of distinction between the male and female cards of a matched pair is recorded in Table A6.15/1, which shows the responses of the whole sample to every pair of cards, and also the median value associated with each card.

When the teachers' responses to the 12 card pairs were analysed all together, rather than as separate pairs, then it was found that the male card of a matched pair evoked higher expectations than the female card in

Table A6.15/2 Cases of sex bias in teachers' expectations

	Magnitude of rating difference		
Sex favoured	1 scale point	2 scale points	3 and more scale points
Boy	54	8	11
Girl	43	13	2

$\chi^2 = 7.05$ $p < 0.05$

Table A6.15/3 Allocation of boys/girls to exam/non-exam conditions

(A) FREQUENCIES

		Girl	
		Pass	Fail
Boy	Pass	202	19
	Fail	13	186

(B) PROPORTIONS

		Girl	
		Pass	Fail
Boy	Pass	0.481	0.045
	Fail	0.031	0.443

$z = 1.037$

The difference between the two correlated proportions is insignificant

Table A6.15/4 Sex split between rating positions 5 and 6

(A) FREQUENCIES

		Girl	
		5 & higher	6
Boy	5 & higher	319	26
	6	15	60

(B) PROPORTIONS

		Girl	
		5 & higher	6
Boy	5 & higher	0.759	0.062
	6	0.036	0.143

$z = 1.70$

The difference between the two correlated proportions is significant at the 5% level (One-tailed test)

17.4% of the 420 comparisons made, whilst the female card produced higher expectations in only 13.8% of the comparisons. Furthermore, when male and female cards were evaluated differently, the gap between their ratings was significantly greater when the male card was favoured than when the female card was favoured (see Table A6.15/2).

Further analysis revealed that the allocation of boys and girls to GCE examination or non-examination groups was not significantly different (see Table A6.15/3). However, the teachers did display sex bias in their expectations for the weakest pupils. They were more emphatic that weak girls should definitely not follow O level courses than they were about weak boys. This finding emerged from comparing the number of boys and girls given the lowest rating (see Table A6.15/4).

The importance of the different factors, besides pupil sex, to teachers' expectations was investigated by ranking the cards on the basis of the number of times each card was placed in the pile that signified O level suitability. These results are reported in Appendix 6.16. The cards were also ranked separately for each sex, but the two sequences were not appreciably different. This suggests that none of the other variables investigated were interacting with pupil sex in any consistent manner to influence teachers' expectations.

6.15.2.A Discussion

Comparing the results of the Marking Exercise, in which teachers' marking patterns were influenced by pupil sex, with the results of the Cards Exercise, in which teachers' judgements were almost unbiased by pupil sex, it would appear that teachers only form sex differentiated expectations for pupils when they are supplied with sufficient information about the pupils. In the present research, samples of pupils' work constituted sufficient information to bring about biasing effects.

A number of studies have shown that when information about an imaginary person is scant, subjects make use of stereotyped beliefs to arrive at perceptions and judgements of that person, and that the availability of more information decreases the subjects' reliance upon stereotyped beliefs (Delia, 1972; Locksley et al., 1980; Rosen & Jerdee, 1974). However, like the present research, a few studies have produced the reverse relationship. Darley & Gross (1983) found that very limited information about a stimulus person did not cause the subjects to express stereotyped expectations and judgements, but that more information did. They proposed a two stage expectancy-confirmation process to account for their findings. When information is very limited and not seen as a valid basis for judgements, subjects will resist the temptation to express expectations, perceptions, or judgements solely on the basis of stereotypes. In the absence of additional information or more relevant information, judgements are consistent with normative expectations. However, when more information is available, even if it is not necessarily very valid information, subjects then make use of that information to confirm their biased expectations. Having satisfied themselves that the evidence supports their expectations, they then express biased expectations and judgements.

The above hypothesis can be used to account for the findings reported in this thesis. When teachers are asked to judge pupils' suitability for O level science courses on the basis of minimal information, they give similar judgements for both boys and girls. Although their latent expectations for boys and for girls may be different, they are not supplied with sufficient information to trigger and/or justify the expression of different expectations. However, when they are supplied with more information in the form of samples of the pupils' work, then as a consequence of their preconceived ideas and beliefs, the teachers perceive and interpret the work differently depending upon whether it was produced by a boy or a girl. In this way the teachers regard the work samples as evidence that both confirms and justifies their sex differentiated expectations. Having convinced themselves that they have made rational judgements, the teachers then express sex differentiated expectations that conform to their latent expectations.

The implications of the present findings to classroom practices are uncertain. The indications are that the availability of more information results in greater bias effects. If this is a stable relationship, then logically science teachers should display even greater bias towards their own pupils, whom they possess a great deal of information about, than they displayed in the Marking Exercise. Many teachers would counter that since they know their pupils well they treat them as individuals. They pay no attention to the sex of a pupil, and thus sex biasing effects could not operate. However, the work of Stanworth (1981) has shown that teachers do tend to group pupils on the basis of their sex, they do tend to value girls less than boys, and they do tend to hold lower expectations for girls than boys. Only detailed research into science teachers operating in natural settings can resolve the extent to which science teachers' expectations and judgements are affected by pupil sex in their day to day teaching.

Table A6.16/1 Rank order of O level suitability

Rank	Number of allocations	Card description
1	69	Very intelligent, must work harder
2	67	High marks, must work harder
3	58	Very intelligent, dislikes science
4	54	Very intelligent, low marks
5	53	High marks, dislikes science
6	44	Average intelligence, high marks
7	29	Likes science, must work harder
8	27	Average intelligence, likes science
9	24	Tries hard, average intelligence
10	10	Tries hard, dislikes science
11	1	Likes science, low marks
12	0	Tries hard, low marks

APPENDIX 6.16

ALLOCATION OF PUPILS TO O LEVEL COURSES

The Cards Exercise produced data that allowed the influence of various pupil characteristics (apart from sex) upon teachers' expectations to be assessed. The importance of different characteristics was compared by ranking the cards on the basis of the number of times each card was placed in the pile that signified O level suitability. Since each card was presented in two formats (boy and girl) to 35 teachers, the total number of times that each basic description could be allocated to the O level pile was 70. The actual number of allocations are recorded in Table A6.16/1.

The table shows that pupils who were described as being very intelligent were most likely to be recommended for an O level course. Of the 210 judgements made about very intelligent pupils (3 basic descriptions x 2 formats x 35 teachers), 181 indicated O level suitability. The next most important factor that influenced teachers' expectations was a pupil's capacity to obtain high marks. Cards describing pupils who generally got high marks were placed in the O level pile 164 times. The influence of the other factors under investigation was considerably less.

The findings clearly indicate that when teachers are asked to recommend fictitious pupils for O level courses, they are greatly influenced by high ability, as signified by either a pupil's intelligence or attainment. However, they are little influenced by pupils' efforts in science or their affective responses to science.

APPENDIX 6.17

THE SCHOOL DETAILS QUESTIONNAIRE

SCHOOL DETAILS

I intend to group teachers according to the type of school in which they are currently teaching. This information can most efficiently be obtained from a single source within each school, so I would be obliged if you would supply the following information for your school. Please answer all the questions as incomplete returns will seriously limit subsequent analysis. All the replies will be treated as strictly confidential. They will be analysed on a group basis and the names of individual schools will not be reported.

1. Type of school:

   (a) Secondary modern ☐
   Grammar ☐
   Comprehensive ☐
   Independent ☐
   Other (specify) ☐ ____________

   (b) Co-educational ☐
   Single sex: boys' ☐
   girls' ☐

   (c) Has the school been in its present form (i.e. as indicated in (a) and (b) above) for at least the past two years? Yes ☐ No ☐

2. Age range of pupils ________

3. Size of school:
   Under 600 pupils ☐
   601 - 1000 ☐
   1001 and over ☐

4. Is the school catchment area mainly
   Large city - inner ☐
   Large city-suburban ☐
   Large town ☐
   Rural ☐

5. Is the background of the pupils generally
   Prosperous ☐
   Average ☐
   Disadvantaged ☐
   Mixed ☐

6. Number of teachers teaching science ________

Table A7.1/1 Male and female non-science teachers' mean ratings of the science subjects

A. Physics

Characteristic	Male (N=128)		Female (N=98)
Practical-theoretical	4.28		4.14
Numerical-verbal	2.11		2.00
Science-arts	1.15		1.09
Logical-intuitive	1.63		1.46
Masculine-feminine	3.25		3.34
Factual-opinionative	1.60	*	1.38
Routine-creative	3.26		2.99
Complex-simple	2.36	**	1.93
Important-unimportant	2.04		2.07

B. Chemistry

Characteristic	Male (N=128)		Female (N=99)
Practical-theoretical	4.06		3.85
Numerical-verbal	2.69		2.69
Science-arts	1.17		1.16
Logical-intuitive	2.01		1.82
Masculine-feminine	3.46	*	3.68
Factual-opinionative	1.52		1.63
Routine-creative	3.15		3.22
Complex-simple	2.27		2.12
Important-unimportant	2.45		2.23

C. Biology

Characteristic	Male (N=128)		Female (N=99)
Practical-theoretical	3.90		3.81
Numerical-verbal	4.23		4.13
Science-arts	1.55		1.56
Logical-intuitive	2.27		2.22
Masculine-feminine	4.34	*	4.14
Factual-opinionative	1.76		1.78
Routine-creative	3.13		3.11
Complex-simple	3.02		2.84
Important-unimportant	2.42	***	1.90

* Significant at 5% level
** Significant at 1% level
*** Significant at 0.1% level

APPENDIX 7.1

FURTHER PERCEPTIONS OF THE CHARACTERISTICS OF SCHOOL SCIENCE

7.1.1.A Responses from male and female non-science teachers

Mean ratings, indicating secondary non-science teachers' perceptions of the science subjects as they are taught in secondary schools up to CSE/O level standard, are presented in Table A7.1/1. The replies received from male and female teachers are shown separately. Inspection of the values indicates that all three science subjects were generally identified more closely with the masculine/science pole of each adjective pair by the female teachers than by the male teachers. The t test (two-tailed) was used to assess the statistical significance of the differences between the mean ratings from men and women. Five of the comparisons were statistically significant at the 5% level or better. Two of these differences arose from the masculine-feminine scale. Male teachers regarded biology to be significantly more feminine, and chemistry to be significantly more masculine than did female teachers. Also of interest were the findings that female teachers considered physics to be significantly more complex and biology to be significantly more important than did male teachers.

7.1.2.A Responses from school teachers teaching at different educational levels

One-way analysis of variance was used to investigate the effect of the level at which teachers work upon their impressions of the school science subjects. The results are recorded in Table A7.1/2. They show that the three groups of teachers were giving significantly different ratings on approximately half of the semantic differential scales. The primary teachers clearly believed the science subjects to be very much more practical orientated than did the secondary teachers. In contrast, on all the other scales on which differences were detected, the primary teachers gave more neutral ratings than the secondary teachers. Thus

Table A7.1/2 Mean ratings awarded to the science subjects by teachers of all levels

A. Physics

	Primary teachers (N=99)	Middle teachers (N=120)	Secondary teachers (N=288)	F	p
Practical-theoretical	3.46	3.87	4.13	7.61	0.001
Numerical-verbal	2.79	2.38	2.04	17.57	0.001
Science-arts	1.49	1.29	1.14	12.66	0.001
Logical-intuitive	1.69	1.70	1.55	1.51	ns
Masculine-feminine	3.57	3.31	3.31	3.08	0.05
Factual-opinionative	1.92	1.76	1.53	7.34	0.001
Routine-creative	3.20	3.18	3.27	0.18	ns
Complex-simple	2.35	2.05	2.15	1.86	ns
Important-unimportant	2.22	2.19	2.00	1.50	ns

B. Chemistry

	Primary teachers (N=101)	Middle teachers (N=120)	Secondary teachers (N=289)	F	p
Practical-theoretical	3.21	3.50	3.90	10.99	0.001
Numerical-verbal	2.87	3.13	2.73	5.17	0.01
Science-arts	1.48	1.31	1.19	7.65	0.001
Logical-intuitive	2.01	2.04	1.91	0.68	ns
Masculine-feminine	3.67	3.59	3.57	0.68	ns
Factual-opinionative	1.95	1.83	1.58	7.31	0.001
Routine-creative	3.20	3.12	3.25	0.33	ns
Complex-simple	2.39	2.47	2.21	2.18	ns
Important-unimportant	2.15	2.33	2.27	0.65	ns

C. Biology

	Primary teachers (N=101)	Middle teachers (N=120)	Secondary teachers (N=289)	F	p
Practical-theoretical	3.30	3.59	3.81	5.88	0.01
Numerical-verbal	3.98	4.38	4.17	2.44	ns
Science-arts	1.88	1.81	1.57	5.46	0.01
Logical-intuitive	2.37	2.31	2.29	0.16	ns
Masculine-feminine	4.07	4.17	4.26	3.98	0.05
Factual-opinionative	1.89	1.95	1.81	0.91	ns
Routine-creative	3.29	3.31	3.24	0.12	ns
Complex-simple	2.86	2.90	2.92	0.07	ns
Important-unimportant	2.27	2.18	2.15	0.34	ns

the science subjects were judged to be more scientific, logical and factual by the secondary teachers than by the primary teachers. Perceptions of the gender connotations of the science subjects are particularly interesting. The secondary teachers tended to exaggerate the gender of all three science subjects, with the result that the secondary teachers rated physics more masculine and biology more feminine than did the primary teachers.

APPENDIX 7.2

DETAILED PERCEPTIONS OF SEX DIFFERENCES IN WRITTEN WORK

Table A7.2/1 Percentage of science teachers (N=89) who believe that they can generally distinguish between the written work of boys and girls (raw cell frequencies in brackets) by teaching subject and teacher's sex

	Men	Women	Total
Physics	95.5 (21)	28.6 (2)	79.3 (23)
Chemistry	88.2 (15)	100 (5)	90.9 (20)
Biology	50.0 (6)	56.3 (9)	53.6 (15)
Integrated Science	62.5 (5)	50.0 (1)	60.0 (6)
Total	79.7 (47)	56.7 (17)	71.9 (64)

Table A7.2/2 Percentage of science teachers (N=89) who can recognise differences between the written work of boys and girls (raw cell frequencies in brackets) by teaching subject and teacher's sex

	Men	Women	Total
Physics	95.5 (21)	57.1 (4)	86.2 (25)
Chemistry	94.1 (16)	100 (5)	95.5 (21)
Biology	50.0 (6)	93.8 (15)	75.0 (21)
Integrated Science	87.5 (7)	100 (2)	90.0 (9)
Total	84.7 (50)	86.7 (26)	85.4 (76)

Table A7.2/3 Features that are viewed as typical of girls' and boys' written work by science teachers

(Of those teachers who listed differences between the written work of boys and girls (N=76), the percentage mentioning the different features in a complimentary (+) or critical (-) manner is recorded)

	Girls		Boys	
	+	-	+	-
Handwriting				
Rounded	5.3	-	-	-
Regularity	6.6	-	-	-
Neatness	6.6	-	-	-
Legibility	5.3	-	-	-
Appearance				
Neatness	77.6	-	-	65.8
Presentation	30.3	-	-	15.8
Diagrams	15.8	-	-	5.3
Details	6.6	-	-	-
Approach				
Conscientiousness	18.4	-	-	11.8
Haste	-	-	-	9.2
Content - Language	-	-	-	5.3
Content - Appraisal				
Thoroughness	18.4	-	-	13.2
Accuracy	9.2	9.2	15.8	-
Understanding	-	10.5	9.2	-
Content - Quantity				
Lengthy	21.1	-	-	-
Brief	-	-	21.1	-

Table A7.2/4 Features that are viewed as typical of girls' and boys' written work by non-science teachers

(Of those teachers who listed differences between the written work of boys and girls (N=83), the percentage mentioning the different features in a complimentary (+) or critical (-) manner is recorded)

	Girls		Boys	
	+	-	+	-
Handwriting				
Large	4.8	-	-	-
Upright	4.8	-	-	-
Rounded	6.0	-	-	-
Regularity	4.8	-	-	-
Neatness	14.5	-	-	-
Legibility	6.0			
Appearance				
Neatness	71.1	-	-	44.6
Presentation	26.5	-	-	20.5
Diagrams	7.2	-	-	-
Details	10.8	-	-	-
Approach				
Concientiousness	21.7	-	-	19.3
Content - Language	8.4	-	-	9.6
Content - Style				
Objective	-	-	8.4	-
Expressive	4.8	-	-	-
Imaginative	4.8	4.8	7.2	-
Content - Appraisal				
Thoroughness	-	-	-	4.8
Understanding	-	-	7.2	-
Content - Quantity				
Lengthy	13.3	-	-	-
Brief	-	-	12.0	-

Table A7.3/1 Secondary school teachers mean ratings of pupils' preferences for subject characteristics (Sample M)

A. Girls' preferences

Characteristic	Science teachers (N=52)		Non-science teachers (N=222)
Practical-theoretical	3.88	*	3.43
Numerical-verbal	5.17		5.22
Science-arts	4.49	(+)	4.96
Logical-intuitive	4.63		4.50
Masculine-feminine	5.19		5.16
Factual-opinionative	4.13		4.50
Routine-creative	4.50	(+)	4.89
Complex-simple	4.48		4.42
Important-unimportant	2.98		2.69

B. Boys' preferences

Characteristic	Science teachers (N=52)		Non-science teachers (N=222)
Practical-theoretical	2.83		3.10
Numerical-verbal	3.10		3.30
Science-arts	2.77	(+)	3.04
Logical-intuitive	2.88		3.08
Masculine-feminine	2.54		2.49
Factual-opinionative	3.04		3.32
Routine-creative	4.25		4.27
Complex-simple	4.08		4.05
Important-unimportant	2.50		2.44

* Significant at 5% level

(+) Significant at 10% level

APPENDIX 7.3

FURTHER VIEWS ABOUT PUPILS' PREFERENCES FOR SUBJECT CHARACTERISTICS

7.3.1.A Responses from secondary teachers of different subject areas

Secondary teachers' views regarding pupils' preferences for a number of subject characteristics are presented in Table A7.3/1. There appears to be much agreement among secondary teachers over the characteristics that boys and girls prefer subjects to display. However, there are indications that non-science teachers are more convinced than science teachers that girls prefer practical and creative subjects. Science teachers seem to be more convinced than non-science teachers that both boys and girls prefer science subjects.

7.3.2.A Responses from school teachers teaching at different educational levels

One-way analysis of variance was used to investigate the effect of the level at which a teacher works upon their perception of the subject characteristics preferred by a typical 14-year-old girl and by a typical 14-year-old boy. The results are recorded in Table A7.3/2. They show that primary and middle school teachers generally gave very similar ratings, especially with respect to the subject characteristics preferred by a girl. Although the responses of the secondary teachers are broadly similar to those given by the primary and middle school teachers, there are several notable differences. Secondary teachers believe more firmly that girls prefer subjects that are simple and verbal. Their beliefs about boys' and girls' preferences regarding the gender image of subjects are even more stereotyped. Not only are secondary teachers more extreme in their view that girls prefer feminine subjects, but they also believe more firmly that boys prefer masculine subjects.

Table A7.3/2 Mean ratings of pupils' preferences for subject characteristics by school teachers of all levels

A. Girls' preferences

Characteristic	Primary teachers (N=101)	Middle teachers (N=120)	Secondary teachers (N=280)	F	p
Practical-theoretical	3.74	3.72	3.53	1.73	ns
Numerical-verbal	4.80	4.83	5.20	9.70	0.001
Science-arts	4.91	4.91	4.92	0.01	ns
Logical-intuitive	4.71	4.63	4.54	0.93	ns
Masculine-feminine	4.85	4.79	5.16	5.68	0.010
Factual-opinionative	4.24	4.40	4.10	2.46	ns
Routine-creative	4.97	4.68	4.81	4.36	ns
Complex-simple	4.13	4.13	4.14	4.36	0.050
Important-unimportant	3.01	3.21	2.75	5.71	0.010

B. Boys' preferences

Characteristic	Primary teachers (N=101)	Middle teachers (N=120)	Secondary teachers (N=280)	F	p
Practical-theoretical	3.29	3.14	3.05	1.43	ns
Numerical-verbal	3.10	3.35	3.26	1.38	ns
Science-arts	3.02	3.06	3.00	0.15	ns
Logical-intuitive	3.04	3.19	3.04	0.99	ns
Masculine-feminine	2.78	2.80	2.51	3.25	0.050
Factual-opinionative	3.21	3.47	3.26	1.70	ns
Routine-creative	3.78	4.06	4.26	5.73	0.010
Complex-simple	3.83	3.93	4.06	1.66	ns
Important-unimportant	3.02	3.06	2.47	13.53	0.001

7.3.3.A Responses of teachers and pupils compared

The Preference for Subject Characteristics scale, in a modified form, was completed by a number of secondary school pupils from comprehensive schools. Even though administration of the scale was restricted to the top 50% of the ability range in each school, it was still thought to be advisable to simplify some of the adjective pairs. Thus numerical-verbal, logical-intuitive, factual-opinionative and complex-simple were changed to based on numbers-based on words, involves thought-involves feeling, based on facts-based on opinions and complicated-simple. In addition, pilot work indicated that science-non-science was more meaningful than science-arts. A discussion of the consequences of altering the adjective pairs, their polarity and their order is presented in Appendix 7.4.

Table A7.3/3 presents girls' preferences for different subject characteristics as reported by a group of 14-year-old boys, a group of 14-year-old girls, the science teachers of the girls (sample TSCH), and the science teachers contacted in the main study (sample P). The table clearly shows that the science teachers overrated the attraction of the feminine/arts associated pole of each adjective pair, with the exception of the routine-creative item. Thus girls actually prefer subjects that are more practical, numerical, science, logical, masculine, factual, complex, important and creative than science teachers realize. The boys also tended to give stereotyped responses when asked what characteristics of subjects appeal to girls. In fact, the boys were more extreme than any other group of respondents in their belief that girls like arts subjects which are intuitive, opinionative, simple and not very important.

Table A7.3/4 presents boys' preferences for different subject characteristics as reported by a group of 14-year-old girls, a group of 14-year-old boys, the science teachers of the boys (sample TSCH), and the science teachers contacted in the main study (sample P). The table shows that the teachers overrated the attraction of some characteristics

Table A7.3/3 Mean ratings from teachers and pupils of girls' preferences for subject characteristics

Characteristic	Pupils: Girls (N=536)	Pupils: Boys (N=variable)	Teachers: Girls' teachers (N=64)	Teachers: Science teachers (N=159)
Practical-theoretical	2.75	3.59	3.77	3.45
Numerical-verbal	4.94	5.05	5.23	5.24
Science-arts	4.17	5.09	4.69	4.43
Logical-intuitive	3.74	5.71	4.50	4.34
Masculine-feminine	4.61	5.06	5.03	5.04
Factual-opinionative	3.30	4.76	3.48	3.84
Routine-creative	5.61	5.47	4.50	4.48
Complex-simple	4.32	5.28	4.81	4.64
Important-unimportant	1.89	3.71	2.67	2.94

Table A7.3/4 Mean ratings from teachers and pupils of boys' preferences for subject characteristics

Characteristic	Pupils: Boys (N=453)	Pupils: Girls (N=variable)	Teachers: Boys' teachers (N=64)	Teachers: Science teachers (N=159)
Practical-theoretical	2.61	2.02	2.39	2.80
Numerical-verbal	4.18	2.90	3.36	3.64
Science-arts	3.04	2.14	2.70	2.94
Logical-intuitive	2.91	2.48	3.05	3.14
Masculine-feminine	2.70	1.83	2.42	2.43
Factual-opinionative	2.70	1.81	2.83	3.09
Routine-creative	5.44	4.68	4.38	4.17
Complex-simple	3.83	4.44	4.20	4.25
Important-unimportant	1.72	2.14	2.42	2.79

(e.g. numerical, science, masculine, routine), and underrated the attraction of others (e.g. logical, factual, complicated and important). The girls tended to give even more stereotyped responses than the teachers. They were more firmly of the opinion that boys prefer numerical, logical, masculine, factual, science subjects than any other group of respondents. Interestingly, the girls were also more committed to the view that boys like simple subjects.

The most significant pointers to emerge from Tables A7.3/3 and A7.3/4 are

(a) that girls' preferences for subject characteristics are closer to the characteristics that boys prefer than science teachers realize,

(b) that science teachers over-estimate the appeal to boys of the following subject characteristics: science, numerical, masculine, routine. Since all of these characteristics are closely associated with the physical science subjects (see section 7.1.1.1), they probably over-estimate the appeal of the science subjects to boys.

(c) Since science teachers tend to give more extreme ratings to girls' preferences and boys' preferences than pupils do themselves, this means that the teachers believe that there are greater differences between boys' preferences and girls' preferences than there actually are.

APPENDIX 7.4

STABILITY OF THE SEMANTIC DIFFERENTIAL

In the present study, it occasionally seemed desirable to compare results obtained from non-identical semantic differentials. Most instances involved the various forms of the semantic differential used to determine school subject characteristics. The practice of comparing results obtained from different instruments would normally constitute a serious threat to the internal validity of any ensuing findings. However, in the present circumstances most of the objections to the practice can be satisfactorily countered. Thus it is contended that the remaining reservations are insufficient to prevent comparisons being made. On these grounds it was decided that comparisons could legitimately be made.

The arguments used to justify the decision to compare data from non-identical instruments are presented below. Most of the arguments refer to the robustness of semantic differential scales and their tendency to produce stable results regardless of differences in administration details.

1. A subject's response to an individual semantic differential rating scale is independent of their response to any other semantic differential rating scale.

Evidence to support this argument is provided by studies that compared different semantic differential formats. There are three possible ways of presenting the scales on which concepts are to be rated.

(a) Concepts can be presented one at a time, with each concept followed by all of the scales on which it is to be rated.

(b) A concept and one of the scales on which it is to be rated can be presented as a single item. The various concept-scale combinations are arranged one after another in a rotating order.

(c) Scales can be presented one at a time, with each scale followed by all of the concepts which are to be rated on it.

Measurements made using the three formats are very similar (Osgood et al., 1957; Wells & Smith, 1960). This indicates that ratings of concept-scale combinations are not significantly influenced by the context in which the item is presented.

Since evidence suggests that items are rated individually and independently of other items, it follows that items can justifiably be compared on an individual basis. More specifically, the responses obtained from two or more administrations of a particular concept-scale combination can be compared even if

(a) the context in which the item was presented varied, and

(b) the length of the semantic differential in which the item was embedded varied.

Further evidence regarding these arguments is provided by Macourt (1976) who failed to detect any differences in the consistency of responses to items appearing at different points in a semantic differential.

2. The response to a semantic differential rating scale is not affected by the polarity direction of the scale.

Osgood et al. (1957) and Heise (1970) recommend that the directionality of scales should be alternated to counteract response bias tendencies. This recommendation implies that subjects' responses to a particular scale are not influenced by it polarity direction. Experimental work by Macourt (1976) failed to show that the positioning of the positive end of a semantic differential scale affected the consistency of responses.

3. Responses to semantic differential rating scales with different wording but similar meaning can be compared.

It can be argued that for language to be an effective form of communication, there must be general agreement over the meanings of words. If this is so, then items whose meanings can be accepted as similar, even though they are expressed in different words, should evoke equivalent responses. For example, it is contended that the items, factual-opinionative and based on facts-based on opinions, are sufficiently close in meaning to elicit comparable responses.

4. The responses from two or more different samples to a semantic differential can be compared.

Evidence to support this argument is provided by Norman (1959) who found that the correlation between the mean scale values produced by two comparable samples of student respondents was 0.94. Such a high reliability figure indicates that random error momentary deviations of the raters do not totally overshadow variance due to the concept being rated by those particular raters. Thus true-score variance across subject populations, as well as across concepts and scales, should be readily detectable.

Table A7.5/1 Females' Social Roles scale: Mean scores of male and female teachers

(A) FIRST PILOT

Teacher sex	N	Mean score	t	p
Male	35	65.63	-2.29	0.05
Female	10	78.40		

(B) SECOND PILOT

Teacher sex	N	Mean score	t	p
Male	31	68.87	-2.10	0.05
Female	18	76.28		

(C) MAIN STUDY

Teacher sex	N	Mean score	t	p
Male	110	15.92	-5.30	0.001
Female	53	19.57		

Table A7.5/2 Females' Social Roles scale: Mean scores of different age groups

(A) FIRST PILOT

Age	N	Mean score	t	p
39 & under	36	70.39	1.61	ns
40 & over	9	60.78		

(B) SECOND PILOT

Age	N	Mean score	t	p
39 & under	43	73.16	2.52	0.05
40 & over	6	60.33		

(C) MAIN STUDY

Age	N	Mean score	t	p*
39 & under	102	17.58	1.69	0.05
40 & over	43	16.19		

* One-tailed test, testing predicted relationship

APPENDIX 7.5

POSSIBLE INFLUENCES UPON TEACHERS' ATTITUDES TOWARDS WOMEN'S ROLE IN SOCIETY

Replies to individual items of the Attitudes to Females' Social Roles scale, together with the total scores, were analysed by comparing the responses given b known groups of teachers. The only teacher variable to consistently produce significant differences was that of teacher sex. Women gave significantly more liberal replies than men in both of the pilot studies (Table A7.5/1) and in the main study (see section 7.2.3.2).

The other comparisons investigated generally failed to produce differences that were statistically significant. However, distinct trends that were repeated over most, if not all, of the items emerged for some teacher variables. Since these trends could be indicative of consistent underlying relationships, they are reported here. Differences that reached statistical significance are shown; all other differences are non-significant.

A teacher's age seems to influence their attitude towards women's role in society. Younger teachers gave more liberal replies than older teachers in both of the pilot studies and in the main study (Table A7.5/2). The pilot studies also suggested that biology teachers gave more liberal replies than physics and chemistry teachers (Table A7.5/3). However, this relationship was probably confounded by the variable teacher sex, since a higher proportion of biology teachers are women and women give more liberal replies than men. The data from the main study failed to reproduce the relationship (Table A7.5/3).

Several interesting trends emerged from teacher backgound variables. Teachers from a middle-class background gave more liberal replies than teachers from a working-class background (Table A7.5/4), and respondents whose mothers had only been housewives gave more liberal replies than

Table A7.5/3 Females' Social Roles scale: Mean scores of teachers of different subjects

(A) FIRST PILOT

Subject	N	Mean score
Physics	13	68.00
Chemistry	11	58.91
Biology	17	76.65
Integrated Science	4	61.50

(B) SECOND PILOT

Subject	N	Mean score
Physics	17	70.18
Chemistry	13	70.85
Biology	14	74.43
Integrated Science	5	70.40

(C) MAIN STUDY

Subject	N	Mean score
Physics	40	17.88
Chemistry	56	16.68
Biology	51	16.73
Integrated Science	14	17.64

Table A7.5/4 Females' Social Roles scale: Item means of teachers from different social class backgrounds, main study

Item No.	Working-class (N=72)	Middle-class (N=73)
1	3.56	3.77
2	3.36	3.33
3	3.69	3.82
4	2.85	3.11
5	3.29	3.49

respondents whose mothers had been employed (Table A7.5/5). Both male and female teachers who had received all or part of their own education at a single sex school gave more liberal replies than teachers who had only been educated at coeducational schools (Table A7.5/6).

Certain features of the schools in which the teachers taught can also be linked with teachers' scores. There was a tendency for city teachers to give more liberal replies than rural teachers (Table A7.5/7). And teachers in prosperous areas tended to give more liberal replies than teachers in average or disadvantaged areas (Table A7.5/8).

Table A7.5/5 Females' Social Roles scale: Item means of teachers having housewife or working mothers, main study

Item No.	Housewife mother (N=101)	Employed mother (N=42)
1	3.69	3.52
2	3.33	3.36
3	3.82	3.64
4	3.01	2.86
5	3.47	3.14

Table A7.5/6 Females' Social Roles scale: Item means of teachers who did and did not attend single sex schools, main study

(A) MALE TEACHERS

Item No.	Coeducational education (N=40)		Single sex education (N=55)
1	3.28		3.36
2	3.08		3.36
3	3.35		3.60
4	2.65		2.96
5	2.58	*	3.25

(A) FEMALE TEACHERS

Item No.	Coeducational education (N=16)		Single sex education (N=29)
1	4.31		4.31
2	3.38		3.59
3	4.19		4.45
4	2.81	*	3.52
5	3.88		4.34

* Significant at the 5% level

Table A7.5/7 Females' Social Roles scale: Item means of teachers from schools of different locations, main study

Item No.	Location of school: Inner city (N=23)	Suburban (N=36)	Town (N=65)	Rural (N=34)
1	4.04	3.72	3.60	3.47
2	3.43	3.31	3.26	3.41
3	4.17	3.81	3.68	3.53
4	3.43	2.89	2.88	3.03
5	3.74	3.56	3.28	3.15

Table A7.5/8 Females' Social Roles scale: Item means of teachers teaching children of different backgrounds, main study

Item No.	Background of pupils: Prosperous (N=23)	Average & Mixed (N=126)	Disadvantaged (N=11)
1	3.70	3.67	3.27
2	3.48	3.33	3.18
3	3.57	3.82	3.18
4	3.30	2.94	3.09
5	3.57	3.36	3.27

APPENDIX 7.6

EFFECT SIZE

A statistically significant result allows the null hypothesis to be rejected and so implies that the association under investigation exists. However, statistical significance gives no indication of the strength of the association. For example, whether a t value reaches a level which indicates a significant difference between two means, is partly determined by the size of the samples. So t values cannot be used as measures of the magnitude of differences. When comparisons are to be made between different variables, actual numerical differences between means do not provide satisfactory measures of the magnitude of differences either, because variables are often measured on different scales. A 'standard' index for gauging the magnitude of a phenomenon is required.

The parameter 'effect size' (ES) is increasingly used as an index of the dgree to which an association is present, or the degree to which the null hypothesis is false. A number of measures of effect size are available (e.g. Craig et al., 1976; Fleiss, 1969; Friedman, 1968; Hays, 1974; Smail & Kelly, 1984). An ES measure advocated by Cohen (1977) and called 'd' has been used in this study. It was partly chosen because it has been described in greater detail than some of the other measures. Also the measure has been widely adopted by workers in a variety of fields (Frieze et al., 1982; Glass & Smith, 1978; Smith, 1980), and several workers have written about its interpretation (Cooper, 1981; Rosenthal & Rubin, 1982).

To assist the interpretation of ES indices, sets of conventional values, corresponding to operational definitions of 'small', 'medium' and 'large' effects, have been proposed by Cohen (1977) for each statistical test's ES index. Efforts were made in selecting these operational criteria to use levels of ES which would accord with subjective

Table A7.6/1 Equivalents of d (From Cohen, 1977)

d	U_1	U_2	U_3	r	r^2
0	0.0%	50.0%	50.0%	.000	.000
0.1	7.7	52.0	54.0	.050	.002
0.2	14.7	54.0	57.9	.100	.010
0.3	21.3	56.0	61.8	.148	.022
0.4	27.4	57.9	65.5	.196	.038
0.5	33.0	59.9	69.1	.243	.059
0.6	38.2	61.8	72.6	.287	.083
0.7	43.0	63.7	75.8	.330	.109
0.8	47.4	65.5	78.8	.371	.138
0.9	51.6	67.4	81.6	.410	.168
1.0	55.4	69.1	84.1	.447	.200
1.1	58.9	70.9	86.4	.482	.232
1.2	62.2	72.6	88.5	.514	.265
1.3	65.3	74.2	90.3	.545	.297
1.4	68.1	75.8	91.9	.573	.329
1.5	70.7	77.3	93.3	.600	.360

assessments of average effect sizes such as are encountered in the social sciences. However, it must be remembered that an effect considered to be trivial in some circumstances may be substantial in other circumstances. Cooper (1981) cautions that the evaluation of effect sizes should include considerations of methodology, the general difficulty of explaining the phenomenon, and the importance attached to the detection of any effects.

7.6.1.A The d index

The d index is used when determining the magnitude of differences between means for two independent groups. It is given by the formula:

$$d = \frac{m_A - m_B}{\sigma}$$

where d is the ES index for t tests of means,

m_A and m_B are the means of the two independent groups,

and σ is the standard deviation of the whole sample.

Assuming that the two samples being compared are normally distributed, of equal variability and of equal size, then d can be converted to a number of U values, which measure the percentage non-overlap between the two sample distributions. U_1 measure the percentage of their combined areas that is not overlapping. U_2 measure the percentage of the sample with the larger mean that exceeds the same percentage of the sample with the smaller mean. U_3 measures the percentage of the sample with the smaller mean which is exceed by the upper half of the sample with the larger mean. Values of these three U measures corresponding to d values of 0 - 1.5 appear in Table A7.6/1.

d can also be expressed as r, a correlation coefficient, or as r^2. The latter is particularly useful as it indicates the proportion of the total variance of the dependent variable in the combined samples associated with or accounted for by group membership. Values for r and r^2 also appear in Table A7.6/1.

APPENDIX 7.7

ANALYSIS OF VARIANCE

Essentially, analysis of variance provides an accurate and rapid way of testing whether the means of a variable differ from one group of observations to another. The method identifies and breaks down the variation present in a set of experimental data. The relative magnitude of the variation resulting from the different sources is determined and used to indicate whether a particular part of the variation is greater than expected under the null hypothesis. Analysis of variance yields the statistic F. The significance of F depends not only on the means and standard deviations in the various groups but also on the size of the sample. If the associated probability is less than 0.05, this provides a strong indication that the means of the different groups are not due to a random sampling of people from a homogeneous population, or in other words that some of the differences among the groups are real.

Analysis of variance is closely related to a set of statistical methods known as regression analysis. However, analysis of variance is usually the appropriate method to use when the groups of observations have been determined by an independent variable that was measured at the nominal level. An example of such a nominal level variable employed in this study is teacher's main teaching subject. This variable has three categories resulting in three groups of observations. The dependent variable in an analysis of variance is usually an interval level variable.

Four assumptions underlie the analysis of variance. First, it is assumed that an individual's score is independent of any other score. Second, it is assumed that the variances in the populations from which the samples are drawn are equal (the assumption of homogeneity of variance). Third, the variables in the populations from which the samples are randomly drawn are assumed to have normal distributions. Finally, it is assumed that the effects of various factors on the total variation are additive. Fortunately, minor violations of these assumptions do not greatly affect the results obtained. In particular, quite large departures from the assumptions of normality and homogeneity are required to seriously affect the validity of the inferences drawn from the data (Ferguson, 1976; Iversen & Norpoth, 1976).

In this study, analysis of variance was computed using the ANOVA programme available in the SPSS computer package (Nie et al., 1975). The ONEWAY programme from the same package was used for one-way analysis of variance calculations. ANOVA relies on the general linear hypothesis approach to analysis of variance, i.e. it is basically a stepwise multiple regression method with the necessary dummy variables being created for the user. Thus, it can handle unbalanced designs with unequal cell frequencies.

The F ratios provided by ANOVA are for a fixed-effect model. This model assumes that data is available for all the categories of the explanatory variable(s). Using this model, inferences can only be made to the categories under consideration.

The default option provided by ANOVA for partitioning the sums of squares was selected. This option is called the "classic experimental design" approach. It is the most appropriate approach for analyses in which "the factors do not have a known causal order, but in which main effects may be assumed to be of a higher priority than interaction effects" (Nie et al., 1975, p.408). It was for these reasons that the classic experimental design approach was selected and used in this research.

When reporting analysis of variance results in this thesis, much of the information customarily made available to the reader has frequently been omitted. Most of the tables of results were already large, and it was felt that a lot of extra detail concerning analysis of variance results would make the tables unduly complicated. So to simplify the information contained within tables, and to focus attention upon those main effects and interactions that produced significant results, it was decided to omit details about F values, sums of squares and degrees of freedom from many of the tables.

APPENDIX 7.8

BIAS RESPONDENTS WITH APPROPRIATE TEACHING EXPERIENCE

The majority of the BIAS results are based upon the replies of 306 science teachers who were judged to have appropriate teaching experience for the marking exercise, as they had all taught chemistry and/or integrated science for at least one year. There were 202 men in this sample, 101 women and three respondents of unspecified sex. The respondents taught in a variety of schools, but the majority came from comprehensive schools (83%).

A number of aspects of the previous teaching experience of the sample fitted and qualified the respondents to mark the work samples. Of the 306 teachers, 227 (74%) had taught chemistry for at least one year, and 105 of them (34%) were currently teaching chemistry as their principal teaching subject. Besides being familiar with the subject material under investigation, the majority were also experienced teachers, 76% of them having taught for more than five years (Table A7.8/1). 217 (72%) of the sample were currently teaching pupils of the same age as the pupils who produced the work samples. Furthermore, the teachers' experience of teaching in different types of schools was also apt. 270 (88%) of the teachers had taught in comprehensive schools, and 281 (92%) had taught in coeducational schools at some stage in their teaching career.

Table A7.8/1 Years of teaching experience by appropriately experienced BIAS respondents

Teaching experience (yrs)	No. of teachers	%
Less than 2	18	5.9
2 - 5	56	18.3
5 - 10	87	28.4
10 - 20	106	34.6
Over 20	39	12.7

Table A7.9/1 Effect of order of presentation upon the assessment of experimental write-ups

(A) Work of a high standard

			Assessment											
			Standard				Mark merited				Mark given			
Position	Preceding work	n	$\bar{X}$	sd	F	p	$\bar{X}$	sd	F	p	$\bar{X}$	sd	F	p
1	-	106	2.59	0.60	2.33	0.10	7.23	1.55	3.14	0.05	7.48	1.40	2.39	0.10
2	P	113	2.68	0.54			7.39	1.45			7.65	1.36		
3	A,P	116	2.75	0.47			7.71	1.41			7.88	1.35		

(B) Work of average standard

			Assessment											
			Standard				Mark merited				Mark given			
Position	Preceding work	n	$\bar{X}$	sd	F	p	$\bar{X}$	sd	F	p	$\bar{X}$	sd	F	p
1	-	119	1.83	0.51	1.01	ns	5.47	1.49	2.76	0.10	5.97	1.33	2.46	0.10
2	G	106	1.75	0.58			5.01	1.40			5.58	1.29		
3	P,G	111	1.86	0.55			5.13	1.69			5.72	1.42		

(C) Work of a low standard

			Assessment											
			Standard				Mark merited				Mark given			
Position	Preceding work	n	$\bar{X}$	sd	F	p	$\bar{X}$	sd	F	p	$\bar{X}$	sd	F	p
1	-	112	1.44	0.53	6.04	0.01	3.89	1.50	2.14	ns	5.04	1.39	5.95	0.01
2	A	118	1.28	0.55			3.68	1.69			4.63	1.61		
3	G,A	104	1.20	0.43			3.44	1.59			4.35	1.41		

G - high standard, A - average standard, P - low standard

APPENDIX 7.9

CONTRAST EFFECTS IN THE MARKING EXERCISE

Since the order of presentation of the work samples was varied, marks were obtained for work of three different standards appearing in three different positions. This enabled contrast effects to be investigated. Previous research has shown that the mark awarded to an essay is influenced by the standard of the immediately preceding essays (Daly & Dickson-Markman, 1982; Hales & Tokar, 1975; Hughes et al., 1980).

Table A7.9/1 presents the results obtained in the present study. To simplify the discussion, it was decided to report only data that refers to the global assessment of an experimental write-up. Thus only the overall standard assigned to each piece of work, the mark that it was judged to merit and the mark that it was actually given are recorded in the table. (Standard of work was measured on a 3-point scale, but marks were out of 10)

The assessment made of each of the three experimental write-ups when they appeared first in the marking booklet constitutes a control. The marks awarded reflect the teachers' unbiased judgements. When the work samples appeared in the second or third positions in the booklet, the opportunity existed for contrast effects to operate. By comparing the marks awarded to the write-ups in the three different positions in the booklet, it was possible to show the influence of the standard of preceding samples of work upon work of high, average and low quality.

The results presented in Table A7.9/1 indicate that a good piece of work is assessed more favourably when it follows work of a lower standard than when it is assessed in isolation. Conversely, a poor piece of work is assessed more harshly when it follows work of a higher standard than when it is assessed in isolation. Furthermore, greater contrast effects are produced by two preceeding work samples that are both better or worse than by a single disparate work sample preceding the sample under investigation. One-way analysis of variance revealed that these contrast effects were statistically significant for half of the assessment variables, and that two of the other three variables approached statistical significance.

The influence of the standard of preceding work upon the assessment of a piece of work of average standard is less decisive and more complex. There was a marked tendency for contrast effects to operate in the expected direction, i.e. for work of average standard to be marked lower when it appeared after higher quality work, and for the effect to be lessened when poorer quality work also preceded the average sample of work, but the contrast effects failed to reach statistical significance at the 5% level.

In previous investigations, contrast effects have usually been demonstrated by placing a block of four or five essays of similar high or low standard before a criterion essay of average standard (Daly & Dickson-Markman, 1982; Hales & Tokar, 1975; Hughes et al., 1980). However, this study indicates that as few as one or two pieces of work placed before the criterion work can produce a contrast effect, and bias the marks awarded to the criterion piece of work. It thus follows that the practice commonly adopted by teachers of reading through several pieces of work before commencing to mark a set of work is probably insufficient to prevent contrast effects biasing the marks awarded to the first few pieces of work. If the teacher purposely selects pieces of work that are expected to be very good or very poor, then the contrast effects will be amplified. Another commonly advised tactic of shuffling the order of examination scripts before starting to mark them (Hales & Tokar, 1975) is likely to be equally ineffective.

The work of Hughes et al. (1980) has shown that contrast effects do tend to disappear after a number of essays have been marked. Thus it would appear that the only reliable method of counteracting contrast effects is for a teacher to read a number of pieces of work or examination scripts, before starting to mark the complete set. Unfortunately, we have few indications as to the optimum, or even minimum number of answers that should be read in order to effectively offset contrast effects.

Table A7.10/1 Mean marks awarded to work of differing standards (N=336)

	Standard of work		
	Good	Average	Poor
Marks merited (m)	7.45	5.21	3.68
Marks given (g)	7.68	5.76	4.68
Difference (g - m)	0.23	0.55	1.00

Table A7.10/2 Percentage of teachers overmarking and undermarking work of different standards (N=336)

Marks given - Marks merited	Standard of work		
	Good	Average	Poor
-3		0.3	
-2	0.3	0.3	
-1	5.7	4.8	1.2
0	69.3	51.2	37.9
1	20.5	29.2	30.7
2	3.9	12.2	21.8
3	0.3	0.9	6.9
4		0.9	1.2
5		0.3	0.3

APPENDIX 7.10

EFFECT OF STANDARD UPON MARKS AWARDED

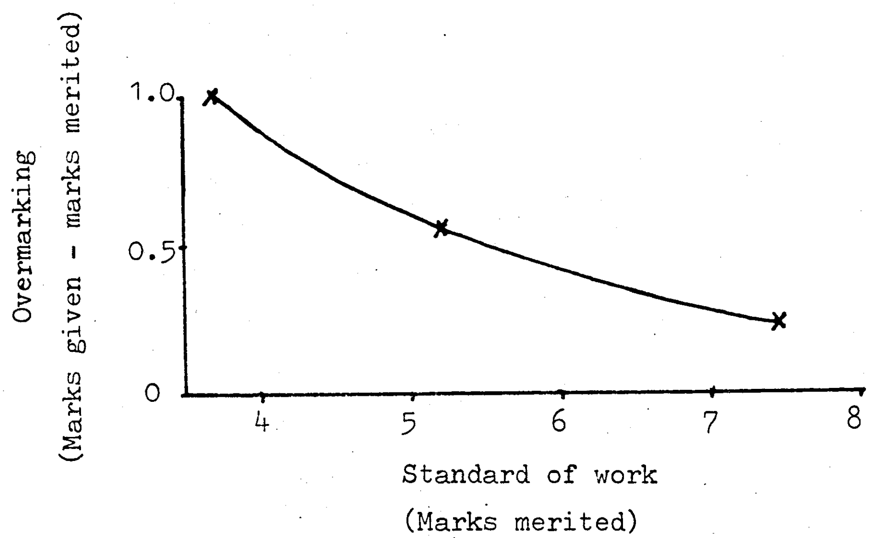
It is logical that work of a low standard should recieve lower marks than work of a higher standard. However, this simple relationship between marks awarded and the standard of a piece of work is complicated by the fact that teachers have a tendency to be generous in their allocation of marks and the degree of generosity is related to the standard of the work. Thus the marks awarded to work of a high standard reflects fairly accurately the true worth of the work, whereas the marks awarded to work of a low standard are inflated and imply that the work is of a considerably higher standard than it actually is, see Table A7.10/1.

The differences between marks merited and marks given recorded in Table A7.10/1 were plotted against marks merited (Figure A7.10/1). The graph suggests that the relationship is not linear and that the degree to which given marks overrate the standard of a piece of work increases disproportionately as the objectively assessed standard of work falls.

Table A7.10/2 reports the percentage of teachers who overmarked and undermarked (compared to their own assessment of true merit) the three sample pairs of different standards. The figures show that only a quarter of the teachers overmarked the good sample pair, just under half (44%) overmarked the average sample pair, and well over half (61%) overmarked the poor sample pair. In addition to the poor sample pair being overmarked by the greatest proportion of teachers, it was also overmarked to the greatest extent. 30% of the teachers overmarked the poor sample pair by more than one mark. Comparative percentages for the average and good sample pairs were 14.3 and 3.9.

Evidence was encountered to suggest that teachers who overmarked believed that such a practice would be of educational or psychological benefit to the pupil involved, e.g. a weak pupil would be better

Figure A7.10/1 The relationship between overmarking and standard of work



motivated by a more promising mark. There is no reason why the teachers should have departed from their normal marking practices in the marking exercise, so teachers who overmarked in the exercise presumably also tend to overmark when marking their own pupils' work. The practice of overmarking raises a number of queries. Does it really achieve the beneficial educational and psychological objectives intended by the teachers, or does it have the opposite effect and cause the pupils to slacken their efforts in the belief that they are producing work of an acceptable standard? Furthermore, does overmarking lead parents to hold false hopes and aspirations for their children that the children may not be capable of realizing? If this does happen, both parents and children could end up bitterly disappointed.

Table A7.11/1 Mean grades awarded to each sample pair for each variable (N=306)

Variable	Good $\bar{x}$	Good (s.d.)	Average $\bar{x}$	Average (s.d.)	Poor $\bar{x}$	Poor (s.d.)
	Sample pair					
Standard	2.70	(0.52)	1.83	(0.54)	1.33	(0.53)
Mark merited	7.50	(1.41)	5.23	(1.53)	3.72	(1.61)
Mark given	7.73	(1.32)	5.79	(1.33)	4.74	(1.48)
Neatness	4.35	(0.70)	2.78	(0.71)	3.27	(0.89)
Effort involved	4.20	(0.75)	3.30	(0.77)	3.41	(0.76)
Grammar and spelling	3.82	(0.82)	3.14	(0.93)	1.63	(0.88)
Scientific accuracy	3.70	(0.91)	2.55	(0.95)	1.63	(0.71)
Understanding of principles	3.85	(0.99)	2.65	(1.06)	1.53	(0.70)
Clarity of explanation	3.27	(1.02)	3.06	(0.92)	2.59	(1.03)
Standard of diagram	4.53	(0.69)	2.71	(0.83)	1.80	(0.79)
Aptitude for science	4.17	(0.81)	2.93	(0.77)	2.11	(0.78)
Attitude towards science	3.75	(0.79)	3.64	(0.81)	3.54	(0.86)
Interest in science	3.72	(0.80)	3.68	(0.86)	3.27	(0.87)
O level suitability	3.73	(0.96)	2.49	(0.96)	1.95	(0.85)
CSE suitability	3.68	(1.36)	3.83	(0.96)	3.33	(1.01)

APPENDIX 7.11

EFFECT OF STANDARD UPON OTHER DEPENDENT AND INDEPENDENT VARIABLES IN THE MARKING EXERCISE

The higher the standard of a piece of work was judged to be, the higher were the mean ratings awarded to it for all the other variables (see Table A7.11/1). The only variables which did not follow this trend were neatness, effort involved, and grammar and spelling.

Besides having an obvious main effect upon the marks awarded to work samples, the standard of a piece of work could also interact with teacher sex or pupil sex. This would mean that the differences in mean scores for one independent variable,i.e. teacher sex or pupil sex, would be larger at some rather than other levels of the second independent variable, i.e. standard of work.

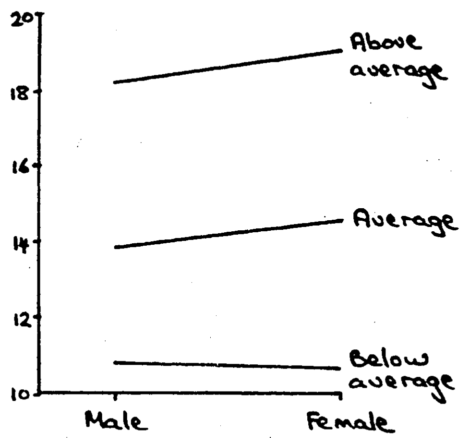
Graphs were drawn to investigate whether first order interactions between standard and pupil sex, and between standard and teacher sex had occurred (Figure A7.11/1). To simplify and magnify any such interactions, the graphs were drawn on the basis of factors, rather than the individual variables which had been used for marking. It had previously been established that there were four common factors or dimensions underlying the set of variables used in the marking exercise (see Appendix 7.12). The first factor referred to the scientific content of the work samples, the second to their presentation, the third to the pupils' affective response to science and the fourth concerned the pupil's potential for science. Mean factor scores were calculated by summing a teacher's or a pupil's ratings for all the variables composing each factor and then calculating the means that male and female teachers awarded, or male and female pupils received, for work of each standard.

The graphs show that mean factor scores awarded to girls for potential were slightly lower than those awarded to boys for all three standards of work. The greatest differences occurred for the affective

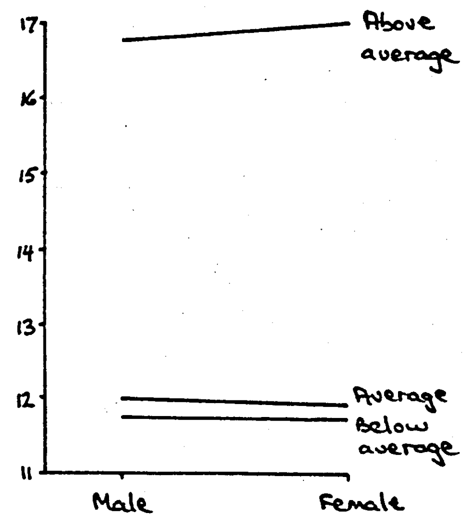
Figure A7.11/1 Effect of standard of work and either teacher sex or pupil sex upon factor scores awarded

(A) Teacher sex and standard of work

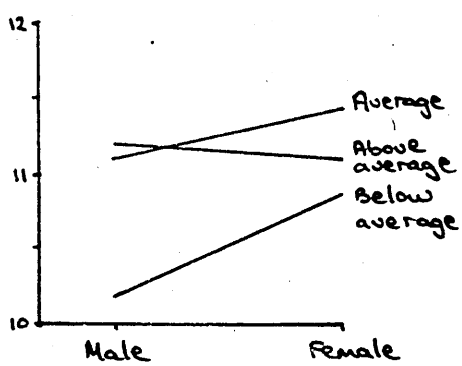
(i) SCIENTIFIC CONTENT



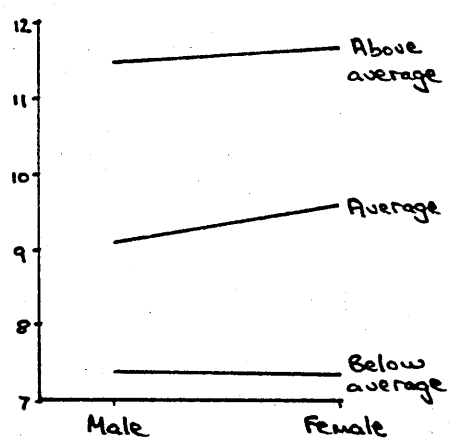
(ii) PRESENTATION



(iii) AFFECTIVE RESPONSE

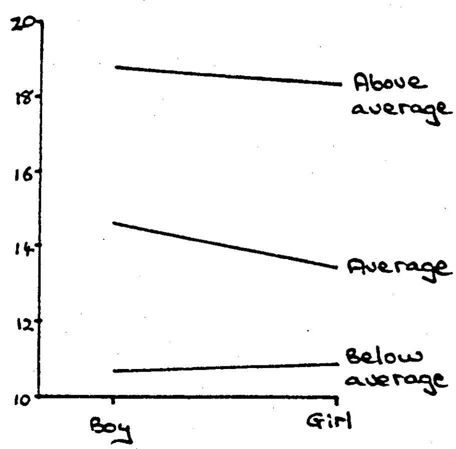


(iv) POTENTIAL

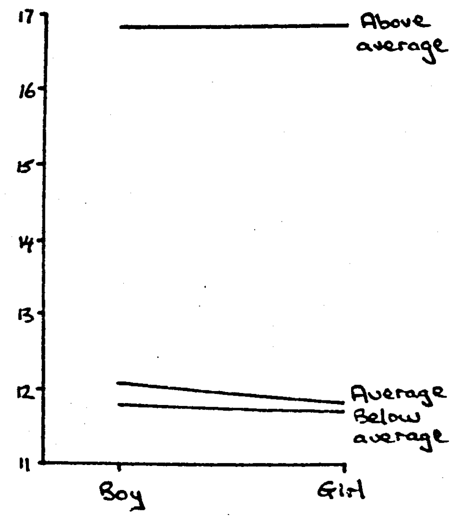


(B) Pupil sex and standard of work

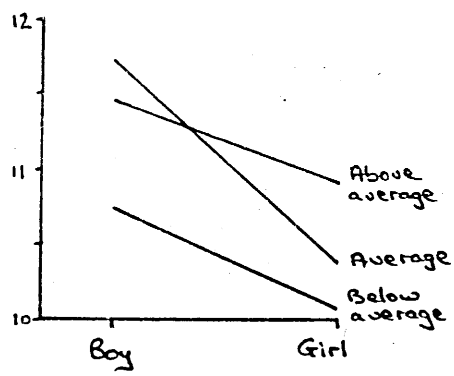
(i) SCIENTIFIC CONTENT



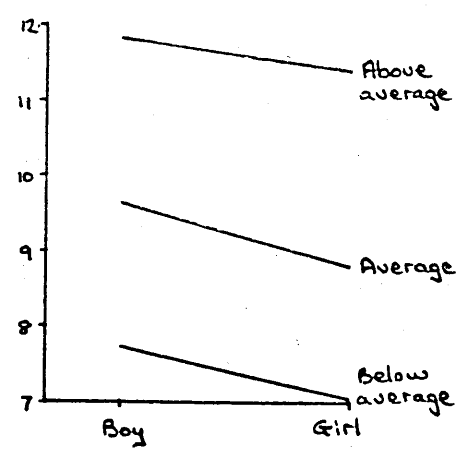
(ii) PRESENTATION



(iii) AFFECTIVE RESPONSE



(iv) POTENTIAL



scores. Female teachers tended to give higher ratings for the affective variables than did the male teachers. Boys consistently received much higher affective scores than girls, regardless of the standard of the work samples.

Most of the curves in most of the graphs are broadly parallel. This implies either that there was no interaction between standard of work and the other two independent variables or that only very weak interactions occurred.

Table A7.12/1 Factor analysis of the ratings awarded by teachers in the Marking Exercise: Varimax factor loadings

Factor	Variable	Rotated factors 1	2	3	4
1	Mark given	49	13	18	29
	Scientific accuracy	86	15	17	07
	Understanding of principles	71	12	01	26
	Clarity of explanation	40	29	22	14
2	Neatness	04	76	03	12
	Effort involved	22	57	25	19
	Grammar & spelling	13	56	20	03
	Standard of diagram	18	52	10	21
3	Attitude towards science	08	20	63	18
	Interest in science	19	14	73	19
4	Aptitude for science	47	21	20	60
	O level suitability	26	20	29	66
	CSE suitability	07	10	13	16

(Decimal points omitted)

APPENDIX 7.12

FACTORS UNDERLYING TEACHERS' MARKING PRACTICES

Replies received to the final form of the BIAS questionnaire were factor analysed to gain insight into the dimensions underlying teachers' marking practices. Only those replies received from teachers with appropriate teaching experience, i.e. those who had taught chemistry and/or integrated science for at least one year (N=306), were included in the analysis. Furthermore, the analysis was restricted to those variables that teachers are most likely to take into account when marking their own pupils' work.

Product moment correlation coefficients between the ratings given to each variable summed across all three sample pairs were calculated. The resulting matrix was factor analysed. The technique of factor analysis is discussed further in Appendix 6.7.

Only the first three factors, accounting for 56.2% of the total variance, had eigenvalues greater than 1. However, since the fourth factor had an eigenvalue of 0.99 and increased the variance accounted for by a further 7.7%, it was decided to also include this factor in the ensuing rotated solution. Rotation was accomplished using the Varimax method. The factor loadings obtained are recorded in Table A7.12/1.

Varimax factors give a more precise identification of the dimensions underlying a set of variables. Careful inspection of the variables that load most highly on each factor permits labels, which attempt to summarize the content of each factor, to be specified and attached to the different dimensions. The factor loadings presented in Table A7.12/1 suggest that Factor 1 should be labelled 'Scientific Content'. Factor 2 contains those variables concerned with the appearance of a piece of work and thus is called 'Presentation'. Factor 3, which consists of the two affective variables, attitude and interest, is simply called 'Affective Response'. Factor 4 contains variables relating to a pupil's future attainment in science and so is labelled 'Potential'.

APPENDIX 7.13

CHI SQUARE

Chi square, symbolized by χ^2, is used when a comparison is made between observed and theoretical frequencies. The theoretical frequencies, or expected frequencies, are produced by some hypothesis which is independent of the experimental data. If the observed frequencies depart significantly from the theoretical frequencies, then the null hypothesis that no difference exists between the observed and theoretical frequencies can be refuted. In addition, the hypothesis or theory that gave rise to the theoretical frequencies can also be rejected.

χ^2 is commonly used in 'tests of goodness of fit' and 'tests of independence'. In both types of tests, observed and expected frequencies are compared.

(A) In tests of goodness of fit, a set of observed frequencies on a single variable is compared with a corresponding set of expected, or theoretical, frequencies. χ^2 is calculated from the formula:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where O = observed frequency

and E = expected frequency.

(B) In tests of independence two variables, usually nominal variables, are involved. The question arises as to whether the two variables are independent of each other, or whether an association exists between them. The data are first arranged in the form of a contingency table. Then the expected frequency of each cell is calculated by multiplying together the totals of the row and column in which the cell is situated and dividing by the overall total.

$$E = \frac{\text{row total x column total}}{\text{grand total}}$$

Once values for the observed and expected frequencies are available, χ^2 can then be calculated as above.

For a 2x2 table, a χ^2 test of independence can be obtained without calculating the expected values. For such a table the formula becomes:

$$\chi^2 = \frac{(bc - ad)^2 k}{efgh}$$

a	c	e
b	d	f
g	h	k

The probability of any value of χ^2 arising by chance is determined from a table of critical values of chi square, entered at the appropriate number of degrees of freedom. The number of degrees of freedom is obtained as the number of classes whose frequency may be assigned arbitrarily. For a contingency table with r rows and c columns, i.e. a r x c table, the number of degrees of freedom is the product (r-1)(c-1).

When expected frequencies are small (less than 10), it is recommended that Yates' correction for continuity be applied (Ferguson, 1976; Lewis, 1967). The obtained frequencies that exceed expectation should be decreased by 0.5, and the obtained frequencies that are less than expectation should be increased by 0.5. This brings the observed and expected values closer together and decreases the value of χ^2. Yates' correction is inappropriate for cases of χ^2 with more than one degree of freedom.

Lastly, χ^2 can only be used appropriately if each and every observation is independent of each and every other observation. Furthermore, each person or reading must occur once and only once in a table.

APPENDIX 7.14

EFFECT OF TEACHER SEX AND PUPIL SEX UPON RATINGS AWARDED IN THE MARKING EXERCISE

The ratings awarded to the three sample pairs were examined separately. Tables A7.14/1 to A7.14/3 present the mean ratings given to male and female pupils by male and female teachers for each variable. These data were submitted to a 2x2 analysis of variance to test the simultaneous effect of pupil sex and teacher sex on the ratings given to each sample pair.

The results of the analysis (included in Tables A7.14/1 to A7.14/3) show that pupil sex produced a number of main effects, especially for understanding of principles, attitude towards science, interest in science, O level suitability and CSE suitability, with a boy being given higher ratings on these variables than a girl. Teacher sex produced fewer main effects, although they were uniformly in the direction of female teachers having awarded higher grades than male teachers. There were very few significant interactions between pupil sex and teacher sex. Where they did occur, examination of the means revealed that it was generally an additive effect, with high marks being received by boys who had been marked by female teachers, and low marks being given to girls who had been marked by male teachers.

Tables A7.14/1 to A7.14/3 also include ES (d) values which indicate the magnitude of the difference between the mean grades awarded to boy' and girls' work by male and female teachers separately. Mean d values for all three sample pairs appear in Table A7.14/4. It can be seen that the mean d values associated with aptitude for science, attitude towards science, interest in science and O level suitability, although mostly small, are significant for both male and female teachers. In contrast, the female teachers alone produced significant d values for standard, marks merited, marks given, scientific accuracy and understanding of principles. The only variable for which male teachers alone produced a significant d value was CSE suitability.

Table A7.14/1 Effect of teacher and pupil sex upon the ratings awarded to the good sample pair

Variable	Teacher sex	Mean grade Boy	Mean grade Girl	d	Analysis of variance Significance Pupil sex	Teacher sex	Inter-action
Standard	Male	2.72	2.65	.13	-	-	-
	Female	2.79	2.68	.22			
Mark merited	Male	7.37	7.44	-.05	-	-	-
	Female	7.98	7.44	.38			
Mark given	Male	7.52	7.79	-.20	-	-	x
	Female	8.14	7.67	.37			
Neatness	Male	4.32	4.31	.01	-	-	-
	Female	4.39	4.40	-.02			
Effort involved	Male	4.11	4.17	-.08	-	-	-
	Female	4.39	4.24	.19			
Grammar and spelling	Male	3.78	3.86	-.10	-	-	-
	Female	3.80	3.84	-.05			
Scientific accuracy	Male	3.70	3.58	.13	-	-	-
	Female	3.98	3.71	.29			
Understanding of principles	Male	3.84	3.63	.21	xx	xxx	-
	Female	4.35	3.91	.46			
Clarity of explanation	Male	3.20	3.27	-.07	-	-	-
	Female	3.46	3.25	.21			
Standard of diagram	Male	4.47	4.53	-.09	-	-	-
	Female	4.63	4.49	.20			
Aptitude for science	Male	4.18	4.05	.16	-	-	-
	Female	4.41	4.18	.30			
Attitude towards science	Male	3.93	3.70	.30	xxx	-	-
	Female	3.91	3.49	.50			
Interest in science	Male	3.89	3.59	.39	xxx	-	-
	Female	3.96	3.51	.52			
O level suitability	Male	3.83	3.66	.19	xx	-	x
	Female	4.04	3.44	.57			
CSE suitibility	Male	3.59	3.73	-.10	-	-	-
	Female	3.61	3.74	-.09			
N	Male	89	113				
	Female	46	55				

Table A7.14/2 Effect of teacher and pupil sex upon the ratings awarded to the average sample pair

Variable	Teacher sex	Mean grade Boy	Mean grade Girl	d	Analysis of variance Significance Pupil sex	Teacher sex	Inter-action
Standard	Male	1.79	1.80	-.02	-	-	-
	Female	1.92	1.87	.10			
Mark merited	Male	5.26	4.85	.26	-	x	-
	Female	5.56	5.49	.05			
Mark given	Male	5.90	5.45	.32	x	-	-
	Female	6.05	5.81	.21			
Neatness	Male	2.82	2.73	.13	-	-	-
	Female	2.83	2.66	.24			
Effort involved	Male	3.28	3.34	-.08	-	-	-
	Female	3.29	3.37	-.11			
Grammar and spelling	Male	3.18	2.94	.27	-	-	-
	Female	3.27	3.24	.03			
Scientific accuracy	Male	2.68	2.33	.37	xx	-	-
	Female	2.63	2.50	.14			
Understanding of principles	Male	2.71	2.35	.35	x	x	-
	Female	2.86	2.79	.06			
Clarity of explanation	Male	3.03	2.92	.12	-	x	-
	Female	3.24	3.21	.03			
Standard of diagram	Male	2.82	2.65	.21	-	-	-
	Female	2.68	2.55	.17			
Aptitude for science	Male	3.03	2.69	.44	x	-	x
	Female	3.02	3.11	-.12			
Attitude towards science	Male	3.74	3.39	.43	xxx	-	-
	Female	3.89	3.45	.56			
Interest in science	Male	3.86	3.27	.67	xxx	-	-
	Female	4.00	3.47	.63			
O level suitability	Male	2.54	2.30	.26	-	-	-
	Female	2.67	2.49	.19			
CSE suitability	Male	4.04	3.44	.63	xxx	-	xx
	Female	3.89	3.89	0			
N	Male	119	83				
	Female	63	38				

Table A7.14/3 Effect of teacher and pupil sex upon the ratings awarded to the poor sample pair

Variable	Teacher sex	Mean grade Boy	Mean grade Girl	d	Analysis of variance Significance: Pupil sex	Teacher sex	Inter-action
Standard	Male	1.34	1.35	-.02	-	-	-
	Female	1.39	1.21	.38			
Mark merited	Male	3.62	3.84	-.13	-	-	-
	Female	3.88	3.55	.21			
Mark given	Male	4.63	3.77	-.09	-	-	-
	Female	4.88	4.81	.05			
Neatness	Male	3.25	3.29	-.04	-	-	-
	Female	3.20	3.33	-.15			
Effort involved	Male	3.45	3.41	.05	-	-	-
	Female	3.29	3.47	-.24			
Grammar and spelling	Male	3.19	3.20	-.01	-	-	-
	Female	3.32	3.28	.05			
Scientific accuracy	Male	1.64	1.67	-.04	-	-	-
	Female	1.73	1.50	.34			
Understanding of principles	Male	1.46	1.69	-.31	-	-	xx
	Female	1.63	1.33	.49			
Clarity of explanation	Male	2.60	2.56	.04	-	-	-
	Female	2.54	2.68	-.14			
Standard of diagram	Male	1.86	1.78	.10	-	-	-
	Female	1.78	1.73	.07			
Aptitude for science	Male	2.23	2.11	.15	-	-	-
	Female	2.15	1.87	.40			
Attitude towards science	Male	3.60	3.36	.28	-	-	-
	Female	3.76	3.57	.22			
Interest in science	Male	3.42	3.00	.49	xxx	x	-
	Female	3.61	3.23	.43			
O level suitability	Male	1.99	1.92	.08	-	-	-
	Female	2.22	1.74	.57			
CSE suitability	Male	3.40	3.09	.31	xx	x	-
	Female	3.70	3.29	.41			
N	Male	103	99				
	Female	41	60				

x Significant at 5% level
xx Significant at 1% level
xxx Significant at 0.1% level
- Not significant

Table A7.14/4 Mean effect sizes (d) of grades awarded to a boy compared to grades awarded to a girl by male and female teachers

Variable	Teacher sex Male	Female
Standard	.03	.23
Mark merited	.03	.21
Mark given	.01	.21
Neatness	.03	.02
Effort involved	-.04	-.05
Grammar and spelling	.05	.01
Scientific accuracy	.15	.26
Understanding of principles	.08	.34
Clarity of explanation	.03	.03
Standard of diagram	.07	.15
Aptitude for science	.25	.19
Attitude towards science	.34	.43
Interest in science	.52	.53
O level suitability	.18	.44
CSE suitability	.28	.11

Table A8.1/1 Mean masculinity ratings given by teachers with working-class and middle-class backgrounds

Subject	Adjective pair	Background Working-class (N=69)		Middle-class (N=72)
Physics	Hard-soft	2.26		2.28
	Tough-tender	2.49		2.78
	Cold-warm	2.75	*	3.21
	Remote-intimate	3.06		3.13
Chemistry	Hard-soft	2.59		2.62
	Tough-tender	2.93		3.15
	Cold-warm	3.64		3.75
	Remote-intimate	3.39		3.35
Biology	Hard-soft	4.30	(+)	3.90
	Tough-tender	4.57		4.30
	Cold-warm	5.09		5.05
	Remote-intimate	4.94		4.84

* Significant at the 5% level
(+) Significant at the 10% level

Table A8.1/2 Mean Masculinity Index scores of teachers with and without experience of teaching in single sex schools

A. Experience of teaching in a boys' school

Subject	Teacher's experience Without experience (N=92)	With experience (N=49)	t	p
Physics	10.86	11.22	-0.62	ns
Chemistry	13.11	11.96	2.13	0.05
Biology	18.85	17.80	1.83	(0.10)

B. Experience of teaching in a girls' school

Subject	Teacher's experience Without experience (N=99)	With experience (N=42)	t	p
Physics	10.82	11.38	-0.92	ns
Chemistry	12.65	12.88	-0.40	ns
Biology	18.81	17.74	1.78	(0.10)

APPENDIX 8.1

ADDITIONAL TEACHER VARIABLES AFFECTING MASCULINITY INDEX SCORES

Marginally significant relationships emerged between Masculinity Index scores and two teacher variables - a teacher's social class during his/her childhood, and his/her experience of teaching in a single sex school. They are reported here since they could be indicative of consistent underlying relationships. Differences that reached statistical significance ($p < 0.05$) or nearly reached statistical significance ($p < 0.10$) are shown. All other differences are non-significant.

Science teachers from a working-class background rated physics more masculine on all four adjective pairs and chemistry more masculine on three of the four adjective pairs than did teachers from a middle-class background (Table A8.1/1). On the other hand, the teachers with a working-class background judged biology to be more feminine on all the four adjective pairs. This suggests that teachers from a working-class background hold more stereotyped views about the gender connotations of the science subjects than do teachers from a middle-calss background. They seem to exaggerate the masculinity of the physical science subjects and the femininity of biology.

Comparing mean Masculinity Index scores given to the three science subjects by teacher who have taught and who have not taught in single sex schools, differences can be detected. One difference reached statistical significance at the 5% level (two-tailed t test), and another two differences approached statistical significance (see Table A8.1/2). There was a tendency for teachers who had taught in boys' schools to view chemistry and biology as more masculine than teachers who had only taught in coeducational schools. Biology also tended to be viewed as more masculine by teachers who had taught in girls' schools.

Table A8.2/1 Effect of a selection of independent and dependent variables upon the proclivity of teachers to favour the work of boys

Variable	N	Xs*	χ^2	p
INDEPENDENT VARIABLES				
Teacher's age				
< 40 years	239	32	0.46	ns
> 40 years	100	29		
Teaching experience				
< 5 years	86	27	3.32	(0.1)
> 5 years	253	35		
Teacher's status				
Ordinary	266	32	0.21	ns
Head of Department	61	30		
Experience of teaching compulsory science				
Has taught	73	34	0.00	ns
Has not taught	251	34		
Experience of teaching in single sex schools				
Has taught in boys' or girls' school	173	33	0.48	ns
No experience	163	30		
Attendance at single sex schools				
Attended boys' or girls' school	208	31	0.22	ns
Attended coeducational school	121	33		
Social class background				
Working-class	160	39	10.60	0.01
Middle-class	175	25		
Mother's occupation				
Housewife	247	33	0.83	ns
Paid employment	85	29		
Sex of current school				
Boys' or girls' school	77	25	3.99	0.05
Coeducational school	260	34		
DEPENDENT VARIABLES				
Chemistry Masculinity Index score				
< 13	62	24	0.00	ns
> 13	63	24		
Composite Masculinity Index score (Biology M.I. - Physics M.I.)				
< 7	38	19	3.61	(0.1)
> 7	73	28		
Females' Social Roles score				
< 17	64	25	0.18	ns
> 17	67	27		
Difference between girls and boys liking of science				
2, 3, 4	34	21	0.71	ns
5	56	25		

* Xs = Number of times the work of a boy received the higher mean rating

APPENDIX 8.2

EFFECT OF INDEPENDENT AND DEPENDENT VARIABLES UPON THE MARKS THAT TEACHERS AWARD TO BOYS AND GIRLS

A very simple measure was chosen to determine the effect of a range of independent and dependent variables upon the marks that teachers award to identical work from boys and girls. For each variable, the sample was dichotomized. Taking teacher status as an example, the sample was split into ordinary teachers and heads of departments. Then the average grades awarded to the three sets of pupil work (3 sets of work x 15 grading variables = 45 grades) by the ordinary teachers when the work was linked with boys' names was compared with the average grades awarded when the work was linked with girls' names. The number of times that higher mean ratings were awarded to the work of a boy was counted and recorded. The procedure was then repeated with the heads of departments marks. Table A8.2/1 shows the number of times different groups of teachers awarded higher mean ratings to the work of a boy. In each case the total number of ratings made was 45. χ^2 was used to assess whether each pair of complementary groups tended to favour the work of boys to differing extents. The χ^2 values and their levels of significance are recorded in Table A8.2/1.

Most of the independent and dependent variables listed in Table A8.2/1 did not influence a teacher's tendency to judge the work of boys more favourably than that of girls. Only two independent variables gave significant results. They were a teacher's social class background and the sex of the school in which he or she was currently teaching. Teachers from working-class backgrounds favoured the work of boys over that of girls to a greater extent than did teachers from middle-class backgrounds. Teachers who were currently teaching in single sex schools favoured the work of boys to a lesser degree than did teachers who were teaching in coeducational schools. One other independent variable produced a marginally significant difference, that of teaching experience. Teachers with more than 5 years teaching experience tended to favour boys' work to a greater extent than did teachers with less teaching experience. Finally, one of the dependent variables also gave a marginally significant difference. Teachers who exaggerated the masculinity of physics and the femininity of biology were slightly more likely to favour boys' work than were teachers who differentiated between the gender connotations of physics and biology to a lesser extent. These results are discussed further and compared with other related findings in Chapters 8 and 9.

REFERENCES

ACE (1979) 'Single sex science' Where, 150, 216.

American Psychological Association (1975) 'Report of the task force on sex bias and sex-role stereotyping in psychotherapeutic practice' American Psychologist, 30, 1169-1175.

Anderson, L.R. and Fishbein, M. (1965) 'Prediction of attitude from the number, strength, and evaluative aspect of beliefs about the attitude object: A comparison of summation and congruity theories' Journal of Personality and Social Psychology, 3, 437-443.

Anderson, N.H. (1961) 'Scales and statistics: Parametric and nonparametric' Psychological Bulletin, 58, 305-316.

Archer, J. (1978) 'Biological explanations of sex-role stereotypes' in J. Chetwynd and O. Hartnett (eds) The sex role system, London, Routledge and Kegan Paul.

Arnold, R.D. (1968) 'The achievement of boys and girls taught by men and women teachers' Elementary School Journal, 68, 367-372.

ASE (1981) Science in Society, London, ASE/Heinemann.

Ashton, B.G. and Meredith, H.M. (1969) 'Attitudes to science and scientists. The attitudes of sixth-formers' School Science Review, 51, 15-19.

Babad, E.Y. (1980) 'Expectancy bias in scoring as a function of ability and ethnic labels' Psychological Reports, 46, 625-626.

Babikian, Y. (1971) 'An empirical investigation to determine the relative effectiveness of discovery, laboratory and expository methods of teaching science concepts' Journal of Research in Science Teaching, 8, 201-209.

Badger, M.E. (1981) 'Why aren't girls better at maths? A review of research' Educational Research, 24, 11-23.

Baker, B.O., Hardyck, C.D. and Petrinovich, L.F. (1966) 'Weak measurements vs strong statistics: and empirical critique of S.S. Stevens' proscriptions on statistics' Educational and Psychological Measurement, 26, 291-309.

Barker Lunn, J.C. (1969) 'The development of scales to measure junior school children's attitudes' British Journal of Educational Psychology, 39, 64-71.

Barker Lunn, J.C. (1970) Streaming in the Primary School, Slough, NFER.

Barker Lunn, J.C. (1972) 'The influence of sex, achievement level and social class on junior school children's attitudes' British Journal of Educational Psychology, 42, 70-74.

Bar-Tal, D. and Frieze, I.H. (1977) 'Achievement motivation for males and females as a determinant of attributions for success and failure' Sex Roles, 3, 301-313.

Bar-Tal, D. and Guttmann, J. (1981) 'A comparison of teachers', pupils' and parents' attributions regarding pupils' academic achievements' British Journal of Educational Psychology, 51, 301-311.

Baruch, G.K. (1972) 'Maternal influences upon college women's attitudes toward women and work' Developmental Psychology, 6, 32-37.

Beardslee, D.C. and O'Dowd, D.D. (1961) 'The college student image of the scientist' Science, 133, 997-1001.

Beckwith, J. and Durkin, J. (1981) 'Girls, boys and math' Science for the People, 13, 6-9, 32-35.

Bendig, A.W. and Hountras, P.T. (1958) 'College student stereotypes of the personality traits of research scientists' Journal of Educational Psychology, 49, 309-314.

Bennett, S.K. (1982) 'Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation' Journal of Educational Psychology, 74, 170-179.

Board of Education (1923) Differentiation of the curriculum for boys and girls respectively in secondary schools, London, HMSO.

Bohrnstedt, G.W. (1970) 'Reliability and validity assessment in attitude measurement' in G.F. Summers (ed) Attitude Measurement, Chicago, Rand McNally.

Bond, L.A. and Deming, S. (1982) 'Children's causal attributions for performance on sex-stereotypic tasks' Sex Roles, 8, 1197-1208.

Boneau, C.A. (1960) 'The effects of violations of assumptions underlying the t test' Psychological Bulletin, 57, 49-64.

Bottomley, J. and Ormerod, M.B. (1981) 'Stability and lability in science interest from middle schools to the age of science choices (14+)' European Journal of Science Education, 3, 329-338.

Bottomley, J. and Ormerod, M.B. (1982) 'Why do children choose chemistry?' Education in Chemistry, 19, 42-43.

Braun, C. (1976) 'Teacher expectation: Sociopsychological dynamics' Review of Educational Research, 46, 185-213.

Bray, J.H. and Howard, G.S. (1980) 'Interaction of teacher and student sex and sex role orientations and student evaluations of college instruction' Contemporary Educational Psychology, 5, 241-248.

Bridgham, R. (1973) 'Grading and enrolments in the sciences' Science Teacher, 40, 41-42.

Brief, A.P. and Wallace, M.J. (1976) 'The impact of employee sex and performance on the allocation of organizational rewards' Journal of Psychology, 92, 25-34.

Briggs, D. (1970) 'The influence of handwriting on assessment' Educational Research, 13, 50-55.

Briggs, D. (1971) 'The handwriting handicap' Where, 58, 170-173.

Briggs, D. (1980) 'A study of the influence of handwriting upon grades using examination scripts' Educational Review, 32, 185-193.

Brophy, J.E. and Good, T.L. (1970) 'Teachers' communication of differential expectations for children's classroom performance: Some behavioral data' Journal of Educational Psychology, 61, 365-374.

Brophy, J.E. and Good, T.L. (1974) Teacher-student relationships: Causes and consequences, New York, Holt, Rinehart and Winston.

Broverman, I.K., Broverman, D.M. and Clarkson, F.E. (1970) 'Sex-role stereotypes and clinical judgments of mental health' Journal of Counseling and Clinical Psychology, 34, 1-7.

Broverman, I.K., Vogel, S.R., Broverman, D.M., Clarkson, F.E. and Rosenkrantz, P.S. (1972) 'Sex-role stereotypes: a current appraisal' Journal of Social Issues, 28, 59-78.

Brown, D.R. (ed) (1968) The role and status of women in the Soviet Union, New York, Teachers College Press.

Brown, M. (1981) 'Christian names in 1980' Letter to The Times, Jan. 5th.

Bruning, J.L. and Albott, W. (1974) 'Funny, you don't look Cecil' in D. Krebs (ed) Readings in Social Psychology: Contemporary Perspectives, New York, Harper and Row, 1976.

BSSRS Women's Collective (1975) 'Power and sex in the laboratory' Science for People, No. 29.

Buchanan, B. and Bruning, J.L. (1971) 'Connotative meaning of names and nicknames on three dimensions' Journal of Social Psychology, 85, 143-144.

Bull, R. and Stevens, J. (1976) 'Do we get 'A' for attractiveness?' Sesame, Dec., p.11.

Bull, R. and Stevens, J. (1979) 'The effects of attractiveness of writer and penmanship on essay grades' Journal of Occupational Psychology, 52, 53-59.

Burger, J.M., Cooper, H.M. and Good, T.L. (1982) 'Teacher attributions of student performance: Effects of outcome' Personality and Social Psychology Bulletin, 8, 685-690.

Burstall, C. (1968) French from eight: A national experiment, Slough, NFER.

Burstall, C. (1978) 'The Matthew effect in the classroom' Educational Research, 21, 19-25.

Busse, T.V. and Seraydarian, L. (1979) 'First names and popularity in grade school children' Psychology in the Schools, 16, 149-153.

Bynner, J.M. (1972) Parents' attitudes to education, London, HMSO.

Cahen, L. (1966) 'An experimental manipulation of the halo effect' Dissertation Abstracts International, 27, 3301A.

Campbell, D.T. (1969) 'Reforms as experiments' American Psychologist, 24, 409-429.

Campbell, D.T. and Fiske, D.W. (1959) 'Convergent and discriminant validation by the multitrait-multimethod matrix' Psychological Bulletin, 56, 81-105.

Campbell, D.T. and Stanley, J.C. (1963) Experimental and quasi-experimental designs for research, Chicago, Rand McNally.

Cannell, C.F. and Kahn, R.L. (1968) 'Interviewing' in G. Lindzey and E. Aronson (eds) Handbook of Social Psychology, Vol. II, Reading, Mass., Addison-Wesley.

Carter, R.S. (1952) 'How valid are marks assigned by teachers?' Journal of Educational Psychology, 43, 218-228.

Carter, R.S. (1953) 'Non-intellectual variables involved in teachers' marks' Journal of Educational Research, 47, 81-95.

Cash, T.F., Gillen, B. and Burns, D.S. (1977) 'Sexism and "beautyism" in personnel consultant decision making' Journal of Applied Psychology, 62, 301-310.

Cattell, R.B. (1952) Factor analysis: An introduction and manual for the psychologist and social scientist, New York, Harper and Row.

Cattell, R.B. and Drevdahl, J.E. (1955) 'A comparison of the personality profile (16PF) of eminent researchers with that of eminent teachers and administrators, and of the general population' British Journal of Psychology, 46, 248-261.

Central Statistical Office (1983) Social Trends, No. 14, 1984 Edition, London, HMSO.

Chase, C.I. (1968) 'The impact of some obvious variables on essay test scores' Journal of Educational Measurement, 5, 315-318.

Chase, C.I. (1979) 'The impact of achievement expectations and handwriting quality on scoring essay tests' Journal of Educational Measurement, 16, 39-42.

Chase, C.I. (1983) 'Essay test scores and reading difficulty' Journal of Educational Measurement, 20, 293-297.

Child, D. (1969) 'A comparative study of personality, intelligence and social class in a technological university' British Journal of Educational Psychology, 39, 40-46.

Chivers, G.E. (1981) Loughborough university women in engineering project. Proceedings of the First GASAT Conference, Eindhoven, Nov.

Chivers, G.E. and Marshall, P. (1983) Attitudes and experiences of some young British women entering engineering education and training courses. Paper presented at Girls and Science and Technology Conference, Oslo, Sept.

Civil Service Department (1979) Civil Service Statistics 1979, London, HMSO.

Claiborn, W.L. (1969) 'Expectancy effects in the classroom: A failure to replicate' Journal of Educational Psychology, 60, 377-383.

Clark, K. (1963) 'Educational stimulation of racially disadvantaged children' in A.H. Passow (ed) Education in depressed areas, New York, Teachers College Press.

Clarricoates, K. (1980) 'The importance of being Ernest ... Emma ... Tom ... Jane. The perception and categorization of gender conformity and gender deviation in primary schools' in R. Deem (ed) Schooling for women's work, London, Routledge and Kegan Paul.

Cline, M.E., Holmes, D.S. and Werner, J.C. (1977) 'Evaluations of the work of men and women as a function of the sex of the judge and type of work' Journal of Applied Social Psychology, 7, 89-93.

Cohen, J. (1977) Statistical power analysis for the behavioral sciences, New York, Academic Press.

Cohen, S.L. and Bunker, K.A. (1975) 'Subtle effects of sex role stereotyping on recruiters' hiring decisions' Journal of Applied Psychology, 60, 566-572.

Collaborative Research (1977) 'Some pupil voices' Collaborative Research Newsletter, 1, 9.

Colman, A.M., Hargreaves, D.J. and Sluckin, W. (1981) 'Preferences for Christian names as a function of their experienced familiarity' British Journal of Social Psychology, 20, 3-5.

Comber, L.C. and Keeves, J.P. (1973) I.E.A. Report, Science education in nineteen countries, New York, John Wiley.

Comrey, A.L. (1978) 'Common methodological problems in factor analytic studies' Journal of Consulting and Clinical Psychology, 46, 648-659.

Cooper, H.M. (1981) 'On the significance of effects and the effects of significance' Journal of Personality and Social Psychology, 41, 1013-1018.

Cooper, H.M. and Burger, J.M. (1980) 'How teachers explain students' academic performance: A catorization of free response academic attribution' American Educational Research Journal, 17, 95-109.

Cowan, R.W. (1971) Schoolchildren's perception of and preference for school subjects, MEd dissertation, University of Edinburgh.

Craig, J.R., Eison, C.L. and Metze, L.P. (1976) 'Significance tests and their interpretation: An example utilizing published research and ω^2' Bulletin of the Psychonomic Society, 7, 280-282.

Crano, W.D. and Mellon, P.M. (1978) 'Causal influence of teachers' expectations on children's academic performance: a cross-lagged panel analysis' Journal of Educational Psychology, 70, 39-49.

Craparo, J.S., Hines, R.P. and Kayson, W.A. (1981) 'Effects of experienced success or failure on self-esteem and problem-solving ability' Psychological Reports, 49, 295-300.

Cronbach, L.J. (1951) 'Coefficient alpha and the internal structure of tests' Psychometrika, 16, 297-334.

Cronbach, L.J. and Meehl, P.E. (1955) 'Construct validity in psychological tests' Psychological Bulletin, 52, 281-302.

Curran, L. (1980) 'Science education: Did she drop out or was she pushed?' in Alice through the microscope by The Brighton Women and Science Group, London, Virago.

Cuttance, P. (1980) 'Do schools consistently influence the performance of their students?' Educational Review, 32, 267-280.

Daly, J.A. and Dickson-Markman, F. (1982) 'Contrast effects in evaluating essays' Journal of Educational Measurement, 19, 309-315.

Dansker, M.M. (1974) 'Counselor evaluation of identical articles attributed to male and female authors' Dissertation Abstracts International, 34, 5621A-5622A.

Darley, J.M. and Gross, P.H. (1983) 'A hypothesis-confirming bias in labeling effects' Journal of Personality and Social Psychology, 44, 20-33.

Datta, L., Schaefer, E. and Davis, M. (1968) 'Sex and scholastic aptitude as variables in teachers' ratings of the adjustment and classroom behavior of Negro and other seventh-grade students' Journal of Educational Psychology, 59, 94-101.

Davidson, H.H. and Lang, G. (1960) 'Children's perceptions of their teachers' feelings toward them related to self-perception, school achievement and behavior' Journal of Experimental Education, 29, 107-118.

Davies, L. and Meighan, R. (1975) 'A review of schooling and sex roles, with particular reference to the experience of girls in secondary schools' Educational Review, 27, 165-178.

Davies, M. and Kandel, D.B. (1981) 'Parental and peer influences on adolescents' educational plans: some further evidence' American Journal of Sociology, 87, 363-387.

Davis, J.A. (1971) Elementary Survey Analysis, Englewood Cliffs, Prentice-Hall.

Dawes, R.M. (1972) Fundamentals of Attitude Measurement, New York, John Wiley.

Deaux, K. (1976) 'Ahh. She was just lucky' Psychology Today, 10, 7, 70-75.

Deaux, K. and Emswiller, T. (1974) 'Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female' Journal of Personality and Social Psychology, 29, 80-85.

Deaux, K. and Farris, E. (1977) 'Attributing causes for one's own performance: The effects of sex, norms and outcome' Journal of Research in Personality, 11, 59-72.

Deaux, K. and Taynor, J. (1973) 'Evaluation of male and female ability: Bias works two ways' Psychological Reports, 32, 261-262.

Delamont, S. (1980) Sex Roles and the School, London, Methuen.

Delamont, S. (1982) 'Sex differences in classroom interaction' EOC Research Bulletin, No. 6, Manchester, Equal Opportunities Commission.

Delia, J.G. (1972) 'Dialects and the effects of stereotypes on interpersonal attraction and cognitive processes in impression formation' The Quarterly Journal of Speech, 58, 285-297.

DES (1980) HMI Series: Matters for Discussion 13, Girls and Science, London, HMSO.

Dickson, C. (1979) 'The state of modern languages' Collaborative Research Newsletter, 5, 3-22.

Douglass, H.R. and Olson, N.E. (1937) 'The relation of high-school marks to sex in four Minnesota senior high schools' School Review, 45, 282-288.

Doyle, W., Hancock, G. and Kifer, E. (1972) 'Teachers' perceptions: Do they make a difference?' Journal of the Association for the Study of Perception, 7, 21-30.

Duckworth, D. and Entwistle, N.J. (1974) 'Attitudes to school subjects: A repertory grid technique' British Journal of Educational Psychology, 44, 76-83.

Duffy, J.C. and Ridinger, B. (1981) 'Stereotyped connotations of masculine and feminine names' Sex Roles, 7, 25-33.

Dunkling, L.A. (1977) First names first, London, Dent.

Dusek, J.B. (1975) 'Do teachers bias children's learning?' Review of Educational Research, 45, 661-684.

Dusek, J.B. and O'Connell, E.J. (1973) 'Teacher expectancy effects on the achievement test performance of elementary school children' Journal of Educational Psychology, 65, 371-377.

Duxbury, J. (1984) 'Girls and physics - the role of a Head of Physics' School Science Review, 65 (No. 233), 648-654.

Dweck, C. (1980) Learned helplessness, or taught helplessness? Paper presented at the Sex Differentiation and Schooling Conference, Cambridge, Jan. 2-5.

Dweck, C. and Bush, E.S. (1976) 'Sex differences in learned helplessness: I Differential debilitation with peer and adult evaluators' Developmental Psychology, 12, 147-156.

Dweck, C., Davidson, W., Nelson, S. and Enna, B. (1978) 'Sex differences in learned helplessness: II The contingencies of evaluative feedback in the classroom and III An experimental analysis' Developmental Psychology, 14, 268-276.

Dweck, C. and Reppucci, N.D. (1973) 'Learned helplessness and reinforcement responsibility in children' Journal of Personality and Social Psychology, 25, 109-116.

Dwyer, C.A. (1979) 'The role of tests and their construction in producing apparent sex-related differences' in M.A. Wittig and A.C. Petersen (eds) Sex-related differences in cognitive functioning, London, Academic Press.

Easlea, B. (1981) Science and Sexual Oppression, Weidenfeld and Nicolson.

Ebbutt, D. (1981a) 'Science options in a girls' grammar school' in A. Kelly (ed) The missing half. Girls and science education, Manchester University Press.

Ebbutt, D. (1981b) 'Girls' science: boys' science revisited' in A. Kelly (ed) The missing half. Girls and science education, Manchester University Press.

Edmiston, R.W. (1943) 'Do teachers show partiality?' Peabody Journal of Education, 20, 234-238.

Edwards, A.L. (1957) Techniques of attitude scale construction, New York, Appleton-Century-Crofts.

Edwards, T.B. and Wilson, A.B. (1958) 'The specialization of interests and academic achievement' Harvard Educational Review, 28, 183-196.

Eggleston, J.F., Galton, M.J. and Jones, M.E. (1976) Processes and products of science teaching, London, Macmillan.

Eisenhardt, W.B. (1977) 'A search for the predominant causal sequence in the interrelationship of interest in academic subjects and academic achievement. A cross-lagged panel correlation study' Dissertation Abstracts International, 37, 4225A.

Elmore, P.B. and LaPointe, K.A. (1974) 'Effects of teacher sex and student sex on the evaluation of college instructors' Journal of Educational Psychology, 66, 386-389.

Engelhart, M.D. (1972) Methods of Educational Research, Chicago, Rand McNally.

EOC (1981) Education of girls - a statistical analysis, Manchester, Equal Opportunities Commission.

EOC (1984) The fact about women is ... 1984 edition, Statistics Unit, Equal Opportunities Commission, Manchester.

Etaugh, C. and Brown, B. (1975) 'Perceiving the causes of success and failure of male and female performers' Developmental Psychology, 11, 103.

Etaugh, C. and Rose, S. (1975) 'Adolescents' sex bias in the evaluation of performance' Developmental Psychology, 11, 663-664.

Etaugh, C. and Sanders, S. (1974) 'Evaluation of performance as a function of status and sex variables' Journal of Social Psychology, 94, 237-241.

Evans, T.D. (1982) 'Being and becoming: Teachers' perceptions of sex-roles and actions toward their male and female pupils' British Journal of Sociology of Education, 3, 127-143.

Fairweather, H. (1976) 'Sex differences in cognition' Cognition, 4, 231-280.

Fawcett Society (1981) 'Science in men's society, Times Educational Supplement, 18.9.1981, p.38.

Feather, N.T. (1966) 'Effects of prior success and failure on expectations of success and subsequent performance' Journal of Personality and Social Psychology, 3, 287-298.

Feather, N.T. (1969) 'Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance' Journal of Personality and Social Psychology, 13, 129-144.

Feather, N.T. (1978a) 'Reactions to male and female success and failure at sex-linked occupations: Effects of sex and socio-economic status of respondents' Australian Journal of Psychology, 30, 21-40.

Feather, N.T. (1978b) 'Perceived affect following male and female success and failure at occupations' Australian Journal of Psychology, 30 133-145.

Feather, N.T. and Simon, J.G. (1971) 'Causal attributions for success and failure in relation to expectations of success based upon selective or manipulative control' Journal of Personality, 37, 527-541.

Feather, N.T. and Simon, J.G. (1975) 'Reactions to male and female success and failure in sex-linked occupations: Impressions of personality, causal attributions, and perceived likelihood of different consequences' Journal of Personality and Social Psychology, 31, 20-31.

Feather, N.T. and Simon, J.G. (1976) 'Stereotypes about male and female success and failure at sex-linked occupations' Journal of Personality, 44, 16-37.

Fee, E. (1982) 'A feminist critique of scientific objectivity' Science for the People, 14, (4) 5-8, 30-33.

Feldman, S.D. (1974) Escape from the doll's house, New York, McGraw-Hill.

Feldman-Summers, S. and Kiesler, S.B. (1974) 'Those who are number two try harder: The effect of sex on attributions of causality' Journal of Personality and Social Psychology, 30, 846-855.

Fennema, E. (1974) 'Mathematics learning and the sexes: A review' Journal for Research into Mathematics Education, 5, 126-129.

Fennema, E. and Sherman, J. (1977) 'Sex-related differences in mathematics achievement, spatial visualization and affective factors, American Educational Research Journal, 14, 51-71.

Ferber, M.A. and Huber, J.A. (1975) 'Sex of student and instructor: A study of student bias' American Journal of Sociology, 80, 949-963.

Ferguson, G.A. (1976) Statistical analysis in psychology and education (4th edition), New York, McGraw-Hill.

Fernberger, S.W. (1948) 'Persistance of stereotypes concerning sex differences' Journal of Abnormal and Social Psychology, 43, 97-101.

Ferry, G. (1982) 'How women figure in science' New Scientist, 1 April, 94, No. 1299, pp. 10-13.

Fielder, W.R., Cohen, R.D. and Feeney, S. (1971) 'An attempt to replicate the teacher expectancy effect' Psychological Reports, 29, 1223-1228.

Finn, J.D. (1972) 'Expectations and the educational environment' Review of Educational Research, 42, 387-410.

Finniston, Sir Montague (1980) Engineering Our Future, Report of the Committee of Inquiry into the Engineering Profession, Cmnd 7794, London, HMSO.

Fishbein, M. and Raven, B.H. (1962) 'The AB scales: An operational definition of belief and attitude' Human Relations, 15, 35-44.

Fisher, R.A. (1950) Statistical methods for research workers (11th edition), New York, Hafner Publishing Company.

Fleiss, J.L. (1969) 'Estimating the magnitude of experimental effects' Psychological Bulletin, 72, 273-276.

Fleming, E.S. and Anttonen, R.G. (1971) 'Teacher expectancy or My Fair Lady' American Educational Research Journal, 8, 241-252.

Fox, L.H., Tobin, D. and Brody, L. (1979) 'Sex role socialization and achievement in mathematics' in M.A. Wittig and A.C. Petersen (eds) Sex-related differences in cognitive functioning, London, Academic Press.

Fransella, F. and Frost, K. (1977) On being a woman, London, Tavistock.

Frazier, N. and Sadker, M. (1973) Sexism in school and society, New York, Harper and Row.

Freedman, S.W. (197) 'How characteristics of student essays influence teachers' evaluations' Journal of Educational Psychology, 71, 328-338.

Friedman, H. (1968) 'Magnitude of experimental effect and a table for its rapid estimation' Psychological Bulletin, 70, 245-251.

Frieze, I.H. (1976) 'Causal attribution and information seeking to explain success and failure' Journal of Research in Personality, 10, 298-305.

Frieze, I.H. (1980) 'Beliefs about success and failure in the classroom' in J.H. McMillan (ed) The social psychology of school learning, New York, Academic Press.

Frieze, I.H. and Snyder, H.N. (1980) 'Children's beliefs about the causes of success and failure in school settings' Journal of Educational Psychology, 72, 186-196.

Frieze, I.H., Whitley, B.E., Hanusa, B.H. and McHugh, M.C. (1982) 'Assessing the theoretical models for sex differences in causal attributions for success and failure' Sex Roles, 8, 333-343.

Gaetano, M.A.K. (1966) 'A study to determine the distribution of male and female figures in elementary science testbooks' Journal of Research in Science Teaching, 4, 178-179.

Gannon, C. (1980) 'Girls underachievement in science' CORE, 4 (1) Microfiche nos 17 & 18.

Garai, J.E. and Scheinfeld, A. (1968) 'Sex differences in mental and behavioral traits' Genetic Psychology Monographs, 77, 169-299.

Gardner, P.L. (1974) 'Sex differences in achievement, attitudes, and personality of science students: A review' Science Education: Research 1974, 231-258.

Garland, H. and Price, K.H. (1977) 'Attitudes toward women in management and attributions for their success and failure in managerial positions' Journal of Applied Psychology, 62, 29-33.

Garman, L.G. and Plant, W.T. (1974) Sex-role stereotypes and educator's descriptions of mature personality. Paper presented at the meeting of the Western Psychological Association, San Francisco, April 27.

Garwood, S.G., Baer, S., Levine, D., Carroll, S. and O'Neal, E. (1981) 'Sex-role expectations of socially desirable first names' Sex Roles, 7, 257-262.

Gitelson, I.B., Petersen, A.C. and Tobin-Richards, M.H. (1982) 'Adolescents' expectancies of success, self-evaluations, and attributions about performance on spatial and verbal tasks' Sex Roles, 8, 411-419.

Glass, G.V. and Smith, M.L. (1978) Meta-analysis of research on the relationship of class-size and achievement, San Francisco, Far West Laboratory.

Goebel B.L. and Cashen, V.M. (1979) 'Age, sex, and attractiveness as factors in student ratings of teachers: A developmental study' Journal of Educational Psychology, 71, 646-653.

Gold, A.R. (1972) 'Reactions to work by authors differing in sex and achievement' Dissertation Abstracts International, 33, 2790B.

Goldberg, P. (1968) 'Are women prejudiced against women?' Transaction, Sept./Oct.

Good, T.L. (1980) 'Classroom expectations: Teacher-pupil interactions' in J.H. McMillan (ed) The social psychology of school learning, New York, Academic Press.

Goodacre, E.J. (1968) Teachers and their pupils' home background, Slough, NFER.

Graf, R.G. and Riddell, J.C. (1972) 'Sex differences in problem-solving as a fuction of problem context' Journal of Educational Research, 65,451-452.

Greenwald, A.G. (1975) 'Consequences of prejudice against the null hypothesis' Psychological Bulletin, 82, 1-20.

Gregory, A.H., Hartley, J.R. and Lewis, D.G. (1969) Basic statistics, London, Methuen Educational.

Griffiths, D. (1976) 'Women and science: A woman replies' in K. Hinton, Women and Science, Siscon Project trial unit, Department of Liberal in Science, University of Manchester.

Griffiths, D. and Saraga, E. (1979) 'Sex differences and cognitive abilities: a sterile field of enquiry?' in O. Hartnett, G. Boden & M. Fuller (eds) Sex-role stereotyping, London, Tavistock Publications.

Gross, M.M. and Geffner, R.A. (1980) 'Are the times changing? An analysis of sex-role prejudice' Sex Roles, 6, 713-722.

Guilford, J.P. (1954) Psychometric Methods, New York, McGraw-Hill.

Guilford, J.P. and Fruchter, B. (1973) Fundamental statistics in psychology and education, New York, McGraw-Hill.

Guttmann, J. and Bar-Tal, D. (1982) 'Stereotypic perceptions of teachers' American Educational Research Journal, 19, 519-528.

Hadley, S.T. (1954) 'A schoolmark - fact or fancy?' Educational Administration and Supervision, 40, 305-312.

Hales, L.W. and Tokar, E. (1975) 'The effect of the quality of preceding responses on the grades assigned to subsequent responses to an essay question' Journal of Educational Measurement, 12, 115-117.

Hall, J. and Frederickson, W.A. (1979) 'Sex role stereotyping, a function of age and education, as measured by a perceptual-projective device' Sex Roles, 5, 77-84.

Hamner, W.C., Kim, J.S., Baird, L. and Bigoness, W.J. (1974) 'Race and sex as determinants of ratings by potential employers in a simulated work-sampling task' Journal of Applied Psychology, 59, 705-711.

Hannon, V. (1981) Ending sex-stereotyping in schools. A sourcebook for school-based teacher workshops, Manchester, Equal Opportunities Commission.

Harari, H. and McDavid, J.W. (1973) 'Name stereotypes and teachers' expectations' Journal of Educational Psychology, 65, 222-225.

Harding, J. (1973) 'What has Nuffield done for girls?' Times Educational Supplement, 30.11.1973, p. 31.

Harding, J. (1979) 'Sex differences in examination performance at 16+' Physics Education, 14, 280-284.

Harding, J. (1981) 'Science in men's society' School Science Review, 63, No. 223, 368-371.

Harding, J. (1982) 'Bring on the girls' Times Educational Supplement, 31.12.1982, p.17.

Harding, J. (1983) 'How the world attracts girls to science' New Scientist, 15 Sept., pp. 754-755.

Harmon, L.W. Birk, J.M., Fitzgerald, L.E. and Tanney, M.F. (eds) (1978) Counseling women, Monterey, Calif., Brooks/Cole.

Harris, M.B. (1975) 'Sex role stereotypes and teacher evaluations' Journal of Educational Psychology, 67, 751-756.

Hart, H. and Olander, E. (1924) 'Sex differences in character as indicated by teachers' ratings' School and Society, 20, 381-382.

Hays, W.L. (1974) Statistics for the social sciences (2nd edition), London, Holt-Saunders.

Hearn, M. (1979) 'Girls for physical science' Education in Science, No. 82, 14-16.

Heikkinen, H. (1978) 'Sex bias in chemistry texts: Where is women's 'place'?' Science Teacher, 45, 16-21.

Heise, D.R. (1969) 'Some methodological issues in semantic differential research' Psychological Bulletin, 72, 406-422.

Heise, D.R. (1970) 'The semantic differential and attitude research' in G.F. Summers (ed) Attitude Measurement, Chicago, Rand McNally.

Henerson, M.E., Morris, L.L. and Fitz-Gibbon, C.T. (1978) How to measure attitudes, Beverly Hills, Sage Publications.

Hesselbart, S. (1977) 'Sex role and occupational stereotypes: Three studies of impression formation' Sex Roles, 3, 409-422.

Hesselbart, S. (1978) 'Racial stereotypes, gender stereotypes and question format' Social Science Quarterly, 59, 338-346.

Hills, P. and Shallis, M. (1975) 'Images of science' New Society, 28 Aug., Vol. 33, 477-478.

Hinton, K. (1976) Women and Science, Siscon Project trial unit, Department of Liberal Studies in Science, University of Manchester.

HMI (1977) Curriculum 11-16, London, DES, HMSO.

HMI (1979) Aspects of secondary education in England, London, HMSO.

Holahan, C.K. and Stephan, C.W. (1981) 'When beauty isn't talent: The influence of physical attractiveness, attitudes toward women, and competence on impression formation' Sex Roles, 7, 867-876.

Hoste, R. (1982) 'Sex differences and similarities in performance in a CSE Biology examination' Educational Studies, 8, (2) 141-153.

Huck, S.W. and Bounds, W.G. (1972) 'Essay grades: An interaction between graders' handwriting clarity and the neatness of examination papers' American Educational Research Journal, 9, 279-283.

Hudson, L. (1966) Contrary imaginations, Harmondsworth, Penguin.

Hudson, L. (1967) 'The stereotypical scientist' Nature, 213, 228-229.

Hudson, L. (1968) Frames of mind, Harmondsworth, Penguin.

Hughes, D.C., Keeling, B. and Tuck, B.F. (1980) 'Essay marking and the context problem' Educational Research, 22, 147-148.

Hughes, D.C., Keeling, B. and Tuck, B.F. (1983) 'Effects of achievement expectations and handwriting quality on scoring essays' Journal of Educational Measurement, 20, 65-70.

Humphreys, L.G. and Stubbs, J. (1977) 'A longitudinal analysis of teacher expectation, student expectation, and student achievement' Journal of Educational Measurement, 14, 261-270.

Hutchings, D. (1967) 'Girls' attitudes to science' New Society, No. 267, 9 Nov., p.671.

Hutt, C. (1979) 'Why do girls underachieve?' Trends, 1979/4.

Iversen, G.R. and Norpoth, H. (1976) Analysis of Variance, Beverly Hills, Sage Publications.

Jackson, P.W., Silberman, M.L. and Wolfson, B.J. (1969) 'Signs of personal involvement in teachers' descriptions of their students' Journal of Educational Psychology, 60, 22-31.

James, A. (1976) 'Does the amount written on assignments bias the grades awarded?' Teaching at a Distance, 7, 49-54.

Jean, P.J. and Reynolds, C. R. (1980) 'Development of the bias in attitudes survey: A sex-role questionnaire' Journal of Psychology, 104, 269-277.

Jenkins, E. (1981) 'Girls and science education - where next?' Education in Chemistry, 18, (2) 37, 59.

Jenkins, J.J., Russell, W.A. and Suci, G.J. (1958) 'An atlas of semantic profiles for 360 words' American Journal of Psychology, 71, 688-699.

Jensen, A. (1969) 'How much can we boost IQ and achievement?' Harvard Educational Review, 39, 1-123.

Jolly, B.C. and Gale, J. (1976) 'Statistical significance: Its appropriate use' Medical Education, 10, 410-411.

José, J. and Cody, J.J. (1971) 'Teacher-pupil interaction as it relates to attempted changes in teacher expectancy of academic ability and achievement' American Educational Research Journal, 8, 39-49.

Kaschak, E. (1978) 'Sex bias in student evaluations of college professors' Psychology of Women Quarterly, 2, 235-242.

Keeves, J. (1973) 'Differences between the sexes in mathematics and science courses' International Review of Education, 19, 47-63.

Kehle, T.J., Bramble, W.J. and Mason, E.J. (1974) 'Teachers' expectations: Ratings of student performance as biased by student characteristics' Journal of Experimental Education, 43, 54-60.

Keller, E.F. (1978) 'Gender and science' Psychoanalysis and Contemporary Thought, 1, 409-433.

Kelly, A. (1976a) 'Women in science: A bibliographic review' Durham Research Review, 7, 1092-1108.

Kelly, A. (1976b) 'Women in physics and physics education' in J.L. Lewis (ed) New trends in physics teaching, Vol. III, Paris, UNESCO.

Kelly, A. (1978a) 'Sex differences in science enrolments: Reasons and remedies' Collaborative Research Newsletter, 4, 61-68.

Kelly, A. (1978b) Girls and science: An international study of sex differences in school science achievement, Stockholm, Almqvist and Wiksell.

Kelly, A. (1979) 'Where have all the women gone?' Physics Bulletin, 30, 108-111.

Kelly, A. (1981) The missing half: Girls and science education, Manchester University Press.

Kelly, A. (1982) 'Why girls don't do science' New Scientist, 20 May, 497-500.

Kelly, A., Alexander, J., Azam, U., Bretherton, C., Burgess, G., Dorney, A., Gold, J., Leahy, C., Sharpley, A. and Spandley, L. (1982) 'Gender roles at home and school' British Journal of Sociology of Education, 3, 281-295.

Kelly, A., Smail, B. and Whyte, J. (1981) Initial GIST survey: Results and implications, Manchester, GIST.

Kelly, A. and Weinreich-Haste, H. (1979) 'Science is for girls?' Women's Studies International Quarterly, 2, 275-293.

Kelly, A., Whyte, J. and Smail, B. (1984) Girls into Science and Technology. Final Report. (Available from GIST, Department of Sociology, University of Manchester)

Keppel, G. (1973) Design and analysis: A researcher's handbook, Englewood Cliffs, N.J., Prentice-Hall.

Kerlinger, F.N. (1972) 'The structure and content of social attitude referents: A preliminary study' Educational and Psychological Measurement, 32, 613-630.

Kerlinger, F.N. (1973) Foundations of Behavioral Research (2nd edition), London, Holt-Saunders.

Keys, W. and Ormerod, M.B. (1977) 'Some sex-related differences in the correlates of subject preference in the middle years of secondary education' Educational Studies, 3, 111-116.

Kim, J. and Mueller, C.W. (1978) Introduction to factor analysis, Quantitative Applications in the Social Sciences, No. 13, Beverly Hills, Sage Publications.

Kirk, R.E. (1968) Experimental design: Procedures for the behavioral sciences, Belmont, CA, Brooks/Cole.

Kitay, P.M. (1940) 'A comparison of the sexes in their attitudes and beliefs about women: A study of prestige groups' Sociometry, 3, 399-407.

Koelsche, C.L. and Newberry, S.L. (1971) 'A study of the relationship between certain variables and the science interests of children' Journal of Research in Science Teaching, 8, 237-241.

Labrant, L.L. (1933) 'A study of certain language developments of children in grades 4-12 inclusive' Genetic Psychology Monographs, 14, 387-491.

Landy, D. and Sigall, H. (1974) 'Beauty is talent: Task evaluation as a function of the performer's physical attractiveness' Journal of Personality and Social Psychology, 29, 299-304.

Lawson, E.D. (1971) 'Semantic differential analysis of men's first names' Journal of Psychology, 78, 229-240.

Leeds Literature Collective (1973) 'Science and girls: A study of primary science textbooks' Shrew, 5, Oct. 1973.

Leirer, V.O., Hamilton, D.L. and Carpenter, S. (1982) 'Common first names as cues for inferences about personality' Personality and Social Psychology Bulletin, 8, 712-718.

Lenney, E. (1981) 'What's fine for the gander isn't always good for the goose: Sex differences in self-confidence as a function of ability area and comparison with others' Sex Roles, 7, 905-924.

Levenson, H., Burford, B., Bonno, B. and Davis, L. (1975) 'Are women still prejudiced against women? A replication and extension of Goldberg's study' Journal of Psychology, 89, 67-71.

Lewis, D.G. (1964) 'The factorial nature of attainment in elementary science' British Journal of Educational Psychology, 34, 1-8.

Lewis, D.G. (1967) Statistical methods in education, University of London Press.

Lewis, I. (1983) 'Some issues arising from an examination of women's experience of university physics' European Journal of Science Education, 5, 185-193.

Lindsay, N.S. (1973) Impact of sex-role stereotypes as reflected in interest in science in the elementary grades, Unpublished EdD dissertation, Harvard University.

Lobban, G. (1974) 'Sex-roles in reading schemes' Forum, 16 (2) 57

Locksley, A., Borgida, E., Brekke, N. and Hepburn, C. (1980) 'Sex stereotypes and social judgment' Journal of Personality and Social Psychology, 39, 821-831.

Loo, R. (1983) 'Caveat on sample sizes in factor analysis' Perceptual and Motor Skills, 56, 371-374.

Lord, F.M. and Novick, M.R. (1968) Statistical theories of mental test scores, Reading, Mass., Addison-Wesley.

Lorenz, J.H. (1982) 'On some psychological aspects of mathematics achievement assessment and classroom interaction' Educational Studies in Mathematics, 13, 1-19.

Lunneborg, P.W. (1970) 'Stereotypic aspect in masculinity-femininity measurement' Journal of Consulting and Clinical Psychology, 34, 113-118.

Maccoby, E.E. (1967) 'Sex differences in intellectual functioning' in E.E. Maccoby (ed) The development of sex differences, London, Tavistock.

Maccoby, E.E. (1976) 'Feminine intellect and the demands of science' in K. Hinton, Women and Science, Siscon Project trial unit, Department of Liberal Studies in Science, University of Manchester.

Maccoby, E.E. and Jacklin, C.N. (1975) The psychology of sex differences, London, Oxford University Press.

MacDonald, M. (1981) 'Schooling and the reproduction of class and gender relations' in R. Dale, G. Esland, R. Fergusson and M. MacDonald (eds) Education and the State, Vol. 2. Politics, Patriarchy and Practice, Lewes, The Falmer Press.

Mackie, M. (1976) 'Students' perceptions of female professors' Journal of Vocational Behavior, 8, 337-348.

Macourt, M.P.A. (1976) An investigation of the consistency of responses to semantic differential scales, North-East Area Study, University of Durham Working Paper 24.

Manstead, A.S.R. and McCulloch, C. (1981) 'Sex-role stereotyping in British television advertisements' British Journal of Social Psychology, 20, 171-180.

Marantz, S.A. and Mansfield, A.F. (1977) 'Maternal employment and the development of sex role stereotyping in 5-11 year old girls' Child Development, 48, 668-673.

Markham, L.R. (1976) 'Influences of handwriting quality on teacher evaluation of written work' American Educational Research Journal, 13, 277-283.

Marshall, J.C. (1967) 'Composition errors and essay examination grades re-examined' American Educational Research Journal, 4, 375-385.

Marshall, J.C. (1972) 'Writing neatness, composition errors, and essay grades re-examined' Journal of Educational Research, 65, 213-215.

Marshall, J.C. and Powers, J.M. (1969) 'Writing neatness, composition errors, and essay grades' Journal of Educational Measurement, 6, 97-101.

Maslin, A. and Davis, J.L. (1975) 'Sex-role stereotyping as a factor in mental health standards among counselors-in-training' Journal of Counseling Psychology, 22, 87-91.

Mason, K.O., Czajka, J.L. and Arber, S. (1976) 'Changes in U.S. women's sex-role attitudes, 1964-1974' American Sociological Review, 41, 573-596.

Masters, R.E. (1978) 'Perceptions of the academic community concerning selected effective teaching behaviors of women professors in masculinely-stereotyped and femininely-stereotyped teaching disciplines' Dissertation Abstracts International, 38 (10-A), 5954.

Matys, G.J. (1970) 'Test and measurement procedures: Review and evaluation' in Examinations at secondary level, Commonwealth Secretariat.

McCandless, B.R., Roberts, A. and Starnes, T. (1972) 'Teachers' marks, achievement test scores, and aptitude relations with respect to social class, race, and sex' Journal of Educational Psychology, 63, 153-159.

McCune, S.D. and Matthews, M. (1975) 'Eliminating sexism: Teacher education and change' Journal of Teacher Education, 26, 294-300.

McDavid, J.W. and Harari, H. (1966) 'Stereotyping of names and popularity in grade school children' Child Development, 37, 453-459.

McDuffie, T.E. (1979) 'Predicting success in ISCS level II' Journal of Research in Science Teaching, 16, 39-44.

McHugh, M.C., Frieze, I.H. and Hanusa, B.H. (1982) 'Attributions and sex differences in achievement: Problems and new perspectives' Sex Roles, 8, 467-479.

McKee, J.P. and Sherriffs, A.C. (1957) 'The differential evaluation of males and females' Journal of Personality, 25, 356-371.

McKennell, A. (1970) 'Attitude measurement: Use of coefficient alpha with cluster or factor analysis' Sociology, 4, 227-245.

McMahan, I.D. (1973) 'Relationships between causal attributions and expectancy of success' Journal of Personality and Social Psychology, 28, 108-114.

McNemar, Q. (1962) Psychological Statistics (3rd edition), New York, John Wiley.

McQuitty, L.L. (1957) 'Elementary linkage analysis for isolating orthogonal and oblique types and typal relevancies' Educational and Psychological Measurement, 17, 207-229.

Mead, M. and Métraux, R. (1957) 'Images of the scientist among high-school students' Science, 126, 384-389.

Measor, L. (1981) Nuffield science and girls. Paper presented at the Sociology of Curriculum Practice Conference, Open University, Sept.

Mendels, G.E. and Flanders, J.P. (1973) 'Teachers' expectations and pupil performance' American Educational Research Journal, 10, 203-212.

Mendelsohn, E. (1976) 'Values and science: A critical reassessment' Science Teacher, 43 (1) 20-23.

Merton, R.K. (1968) Social theory and social structure, New York, Free Press.

Miller, C. and Swift, K. (1979) Words and Women, Harmondsworth, Penguin.

Minden, S. and Duelli-Klein, R. (1982) Review of 'The missing half: Girls and science education' Women's Studies International Forum, 5, 514-516.

Mischel, H.N. (1974) 'Sex bias in the evaluation of professional achievements' Journal of Educational Psychology, 66, 157-166.

Mischel, W. (1967) 'A social-learning view of sex differences in behavior' in E.E. Maccoby (ed) The development of sex differences, London, Tavistock.

Mitroff, I.I., Jacob, T. and Moore, E.T. (1977) 'On the shoulders of the spouses of scientists' Social Studies of Science, 7, 303-327.

Morrell, F. (1981) Reported in Education, 158 (19), 6 Nov., p.349.

Moser, C.A. and Kalton, G. (1971) Survey methods in social investigation, London, Heinemann Educational Books.

Murphy, R.J.L. (1978) 'Sex differences in examination performance: Do these reflect differences in ability or sex-role stereotypes?' Educational Review, 30, 259-263.

Murphy, R.J.L. (1980) 'Sex differences in GCE examination entry statistics and success rates' Educational Studies, 6, 169-178.

Murphy, R.J.L. (1982) 'Sex differences in objective test performance' British Journal of Educational Psychology, 52, 213-219.

Nash, R. (1973) Classrooms observed. The teacher's perception and the pupil's performance, London, Routledge and Kegan Paul.

National Children's Home (1984) Children Today, NCH.

Nelson, S.D. (1977) 'First-name stereotypes and expected academic achievement of students' Psychological Reports, 41, 1343-1344.

Newton, P. (1983) 'Deciding on engineering: Implications of a non-traditional career choice' EOC Research Bulletin, No. 7, Manchester, Equal Opportunities Commission.

Newton, R.F. (1942) 'Do men teachers record higher grades than women teachers?' School and Society, 56, 18 July, p.72.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. and Bent, D.H. (1975) Statistical Package for the Social Sciences (2nd edition), New York, McGraw-Hill.

Nilsen, A.P. (1976) 'Women in children's literature' in M. Guttentag and H. Bray Undoing sex stereotypes, New York, McGraw-Hill.

Norman, W.T. (1959) 'Stability characteristis of the semantic differential' American Journal of Psychology, 72, 581-584.

Nunnally, J.C. (1975) Introduction to statistics for psychology and education, New York, McGraw-Hill.

Nunnally, J.C. (1978) Psychometric Theory (2nd edition), New York, McGraw-Hill.

Nuttall, D.L., Backhouse, J.K. and Willmott, A.S. (1974) Comparability of standards between subjects (Schools Council Examination Bulletin No. 29), London, Evans/Methuen Educational.

O'Connell, E., Dusek, J. and Wheeler, R. (1974) 'A follow-up study of teacher expectancy effects' Journal of Educational Psychology, 66, 325-328.

Oetzel, R.M. (1967) 'Annotated bibliography and classified summary of research in sex differences' in E.E. Maccoby (ed) The development of sex differences, London, Tavistock.

Open University (1973) E341, Methods of educational enquiry: An empirical approach, Milton Keynes, Open University Press.

Open University (1975) D101, Making sense of society, Milton Keynes, Open University Press.

Open University (1976) D305, Social psychology, Milton Keynes, Open University Press.

Open University (1979) DE304, Research methods in education and the social sciences, Milton Keynes, Open University Press.

Open University (1982) D102, Social sciences: A foundation course, Milton Keynes, Open University Press.

Oppenheim, A.N. (1966) Questionnaire design and attitude measurement, London, Heinemann.

Ormerod, M.B. (1975) 'Subject preference and choice in co-educational and single-sex secondary schools' British Journal of Educational Psychology, 45, 257-267.

Ormerod, M.B. (1981) 'Factors differentially affecting the science subject preferences, choices and attitudes of girls and boys' in A. Kelly (ed) The missing half: Girls and science education, Manchester University Press.

Ormerod, M.B., Bottomley, J.M., Keys, W.P. and Wood, C. (1979) 'Girls and physics education' Physics Education, 14, 271-277.

Ormerod, M.B. and Duckworth, D. (1975) Pupils' attitudes to science: A review of research, Slough, NFER.

Orne, M.T. (1962) 'On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications' American Psychologist, 17, 776-783.

Osgood, C.E., Suci, G.J. and Tannenbaum, P.H. (1957) The measurement of meaning, Urbana, University of Illinois Press.

Palardy, J.M. (1969) 'What teachers believe - what children achieve' Elementary School Journal, 69, 370-374.

Palfrey, C.F. (1973) 'Headteachers' expectations and their pupils' self-concepts' Educational Research, 15, 123-127.

Parry, G. (1983) 'A British version of the attitudes towards women scale (AWS-B)' British Journal of Social Psychology, 22, 261-263.

Parsons, J.E., Meece, J.L., Adler, T.F. and Kaczala, C.M. (1982) 'Sex differences in attributions and learned helplessness' Sex Roles, 8, 421-432.

Parten, M. (1950) Surveys, polls, and samples: Practical procedures, New York, Harper.

Payne, S.L. (1951) The art of asking questions, Princeton, New Jersey, Princeton University Press.

Peden, I.C. (1965) 'The missing half of our technical potential: Can we motivate the girls?' Mathematics Teacher, 58, 2

Pedhazur, E.J. (1978) 'Review of the Wilson-Patterson Attitude Inventory' in O.K. Buros (ed) Eighth Mental Measurements Yearbook, Vol. 1, pp. 1150-1152, New Jersey, Gryphon Press.

Penrose, J. (1982) 'The girl who always says yes' The Guardian, 6 April.

Persell, C.H. (1977) Education and inequality. A theoretical and empirical synthesis, New York, The Free Press.

Peterson, K., Kauchak, D. and Yaakobi, D. (1980) 'Science students' role-specific self-concept: course, success, and gender' Science Education, 64, 169-174.

Pheterson, G.I., Kiesler, S.B. and Goldberg, P.A. (1971) 'Evaluation of the performance of women as a function of their sex, achievement, and personal history' Journal of Personality and Social Psychology, 19, 114-118.

Post, R.D. (1981) 'Causal explanations of male and female academic performance as a function of sex-role biases' Sex Roles, 7, 691-698.

Pratt, J. (1984) The attitudes of teachers. Paper presented at the Girl Friendly Schooling Conference, Manchester, Sept. 1984.

Preece, P.F.W. (1977) 'Sex and the physics teacher' Science and Public Policy, 4, 281

Raat, J.H. (1983) Teacher training towards a better girls' science and technology education. Paper presented at the second Girls and Science and Technology Conference, Oslo.

Ray, J.J. and Pratt, G.J. 'Is the influence of acquiescence on "catchphrase" type attitude scale items not so mythical after all?' Australian Journal of Psychology, 31, 73-78.

Reid, M.I., Barnett, B.R. and Rosenberg, H.A. (1974) A matter of choice: A study of guidance and subject options, Windsor, NFER.

Ricks, F.A. and Pyke, S.W. (1973) 'Teacher perceptions and attitudes that foster or maintain sex role differences' Interchange, 4, 26-33.

Rivett, P.J. (1980) The influence of handwriting and tidiness in the assessment of science experiment write-ups, BPhil(Ed) dissertation, University of Birmingham.

Robb, B. and Raven, M. (1981) 'Maternal employment and children's sex-role perceptions' Educational Research, 23, 223-225.

Roe, A. (1952) The making of a scientist, New York, Dodd Mead.

Rosen, B. and Jerdee, T. (1973) 'The influence of sex-role stereotypes on evaluations of male and female supervisory behavior' Journal of Applied Psychology, 57, 44-48.

Rosen, B. and Jerdee, T.H. (1974a) 'Influence of sex role stereotypes on personnel decisions' Journal of Applied Psychology, 59, 9-14.

Rosen, B. and Jerdee, T.H. (1974b) 'Sex stereotyping in the executive suite' Harvard Business Review, 52, March-April, 45-58.

Rosen, H. (1972) Language and class: A critical look at the theories of Basil Bernstein, Bristol, Falling Wall Press.

Rosenkrantz, P., Vogel, S., Bee, H. and Broverman, I. (1968) 'Sex-role stereotypes and self-concepts in college students' Journal of Counseling and Clinical Psychology, 32, 287-295.

Rosenshine, B. (1971) Teaching behaviours and student achievement, Windsor, NFER.

Rosenthal, R. (1973) 'The Pygmalion effect lives' Psychology Today, Sept., pp. 56-63.

Rosenthal, R. (1976) Experimenter effects in behavioral research, New York, Irvington.

Rosenthal, R. and Jacobson, L. (1968) Pygmalion in the classroom, New York, Holt, Rinehart and Winston.

Rosenthal, R. and Lawson, R. (1964) 'A longitudinal study of the effects of experimenter bias on the operant learning of a laboratory rat' Journal of Psychiatric Research, 2, 61-72.

Rosenthal, R. and Rubin, D.B. (1979) 'Comparing significance levels of independent studies' Psychological Bulletin, 86, 1165-1168.

Rosenthal, R. and Rubin, D.B. (1982) 'A simple, general purpose display of magnitude of experimental effect' Journal of Educational Psychology, 74, 166-169.

Ross, C.C. (1947) Measurement in today's schools (2nd edition), New York, Prentice-Hall.

Rotter, J.B. (1966) 'Generalized expectancies for internal versus external control of reinforcement' Psychological Monographs, 80, 1-28.

Rowell, J.A. (1971) 'Sex differences in achievement in science and the expectations of teachers' Australian Journal of Education, 15, 16-29.

Sadker, M. and Sadker, D. (1979) Beyond pictures and pronouns: Sexism in teacher education textbooks. Women's Educational Equity Act Program, U.S. Department of Health, Education, and Welfare.

Sadker, M. and Sadker, D. (1982) Sex Equity Handbook for Schools, New York, Longman.

Samuel, J. (1981) 'Feminism and science teaching: Some classroom observations' in A. Kelly (ed) The missing half: Girls and science education, Manchester University Press.

Saraga, E. (1975) Girls and boys: Are there differences in ability? Paper presented at the Girls and Science Education Conference, Chelsea College, London.

Saraga, E. and Griffiths, D. (1981) 'Biological inevitabilities or political choices? The future for girls in science' in A. Kelly (ed) The missing half: Girls and science education, Manchester University Press.

Sayers, J. (1980) 'Psychological sex differences' in Alice through the microscope by The Brighton Women and Science Group, London, Virago.

Scannell, D.P. and Marshall, J.C. (1966) 'The effect of selected composition errors on grades assigned to essay examinations' American Educational Research Journal, 3, 125-130.

Scarbro, H.A. (1979) 'Sex and locus of control as determinants of children's responses to peer versus adult praise' Journal of Educational Psychology, 71, 604-612.

Schein, V.E. (1973) 'The relationship between sex role stereotypes and requisite management characteristics' Journal of Applied Psychology, 57, 95-100.

Schein, V.E. (1975) 'Relationships between sex role stereotypes and requisite management characteristics among female managers, Journal of Applied Psychology, 60, 340-344.

Schneider, J.F. (1973) 'The Conservatism Scale: Independent of SD?' British Journal of Social and Clinical Psychology, 12, 90-91.

Schrank, W. (1968) 'The labeling effect of ability grouping' Journal of Educational Research, 62, 51-52.

Schrank, W. (1970) 'A further study of the labeling effect of ability grouping' Journal of Educational Research, 63, 358-360.

Scott, M. (1980) 'Teach her a lesson: Sexist curriculum in patriarchal education' in D. Spender and E. Sarah (eds) Learning to Lose, London, The Women's Press.

Seale, C., Bloomfield, J. and Pratt, J. (1982) 'Curricular differences in secondary schools' EOC Research Bulletin, No. 6, Manchester, Equal Opportunities Commission.

Sears, P.S. and Feldman, D.H. (1974) 'Teacher interactions with boys and with girls' in J. Stacey, S. Bereaud and J. Daniels (eds) And Jill came tumbling after: Sexism in American education, New York, Dell.

Selmes, C. (1969) 'Attitudes to science and scientists. The attitudes of 12/13-year-old pupils' School Science Review, 51, 7-14.

Sexton, P.C. (1969) The Feminized Male, New York, Random House.

Sharp, C. and Post, R. (1980) 'Evaluation of male and female applicants for sex-congruent and sex-incongruent jobs' Sex Roles, 6, 391-401.

Sharp, R. and Green, T. (1975) Education and Social Control, London, Routledge and Kegan Paul.

Sharpe, S. (1976) 'Just like a girl' How girls learn to be women, Harmondsworth, Penguin.

Simpson, R.L. (1974) 'Sex stereotypes of secondary school teaching subjects: Male and female status gains and losses' Sociology of Education, 47, 388-398.

Singleton, R. and Christiansen, J.B. (1977) 'The construct validation of a shortform attitude toward feminism scale' Sociology and Social Research, 61, 294-303.

Slade, P. and Jenner, F.A. (1978) 'Questionnaire measuring attitudes to females' social roles' Psychological Reports, 43, 351-354.

Smail, B. and Kelly, A. (1984) 'Sex differences in science and technology among eleven year old schoolchildren: I Cognitive' Research in Science and Technological Education, 2, 61-76.

Smail, B. and Kelly, A. (in press) 'Sex differences in science and technology among eleven year old schoolchildren: II Attitudes' Research in Science and Technological Education.

Smail, B. and Whyte, J. (1983) 'Girls' access to engineering: What can be done in school?' EOC Research Bulletin No. 7, Manchester, Equal Opportunities Commission.

Smail, B., Whyte, J. and Kelly, A. (1982) 'Girls into science and technology: The first two years' School Science Review, 64, No.225, 620-630.

Smead, V.S. and Chase, C.I. (1981) 'Student expectations as they relate to achievement in eighth grade mathematics' Journal of Educational Research, 75, 115-120.

Smith, E.R., Ferree, M.M. and Miller, F.D. (1975) 'A short scale of attitudes toward feminism' Representative Research in Social Psychology, 6, 51-56.

Smith, H.W. (1975) Strategies of social research: The methodological imagination, London, Prentice-Hall.

Smith, M.L. (1980) 'Teacher Expectations' Evaluation in Education, 4, 53-55.

Smithers, A. and Collings, J. (1981) 'Girls studying science in the sixth form' in A. Kelly (ed) The missing half: Girls and science education, Manchester University Press.

Snow, R.E. (1969) 'Unfinished Pygmalion' Contemporary Psychology, 14, 197-199.

Sohn, D. (1982) 'Sex differences in achievement self-attributions: An effect-size analysis' Sex Roles, 8, 345-357.

Soloff, S. (1973) 'Effect of non-content factors on the grading of essays' Graduate Research in Education and Related Disciplines, 6, 44-54.

Solomon, D. and Kendall, A. (1977) 'Dimensions of children's classroom behavior, as perceived by teachers' American Educational Research Journal, 14, 411-421.

Sommer, R. (1965) 'Anchor-effects and the semantic differential' American Journal of Psychology, 78, 317-318.

Spear, M. (1982) Sex bias in a science marking exercise, MEd dissertation, University of Exeter.

Spear, M. (1984) 'The biasing influence of pupil sex in a science marking exercise' Research in Science and Technological Education, 2, 55-60.

Spence, J.T. and Helmreich, R. (1972) The attitudes toward women scale: An objective instrument to measure attitudes toward the rights and roles of women in contemporary society. Journal Supplement Abstract Service, Catalog of Selected Documents in Psychology, 2, 66.

Spence, J.T., Helmreich, R. and Stapp, J. (1973) 'A short version of the Attitudes toward Women Scale (AWS)' Bulletin of the Psychonomic Society, 2, 219-220.

Spender, D. (1977) The facts of life. Sex-differentiated knowledge in the English classroom and the school. Unpublished paper.

Spender, D. (1980) Man Made Language, London, Routledge and Kegan Paul.

Spender, D. (1982) Invisible women: The schooling scandal, London, Writers and Readers Publishing Cooperative.

Stanley, J.C. (1967) 'Further evidence via the analysis of variance that women are more predictable academically than men' Ontario Journal of Educational Research, 10, 49-56.

Stanworth, M. (1981) Gender and schooling: A study of sexual divisions in the classroom, London, Women's Research and Resources Centre.

Stewart, A.J. and Winter, D.G. (1974) 'Self-definition and social definition in women' Journal of Personality, 42, 238-259.

Stiver, S.B. (1976) 'Fear of success: A motive to avoid success or a response to violation of sex-role prescriptions?' Dissertation Abstracts International, 37, 1885B.

Stones, R. (1983) 'Pour out the cocoa, Janet': Sexism in children's books, London, Longman for Schools Council.

Sutherland, M.B. (1981) Sex Bias in Education, Oxford, Blackwell.

Sweeney, P.D., Moreland, R.L. and Gruber, K.L. (1982) 'Gender differences in performance attributions: Students' explanations for personal success or failure' Sex Roles, 8, 359-373.

Tait, K., Hartley, J.R. and Anderson, R.C. (1973) 'Feedback procedures in computer-assisted arithmetic instruction' British Journal of Educational Psychology, 43, 161-171.

Taylor, J. (1979) 'Sexist bias in physics textbooks' Physics Education, 14, 277-280.

Taynor J. and Deaux, K. (1973) 'When women are more deserving than men: Equity, attribution and perceived sex differences' Journal of Personality and Social Psychology, 28, 360-367.

Taynor, J. and Deaux, K. (1975) 'Equity and perceived sex differences: Role behavior as defined by the task, the mode, and the actor' Journal of Personality and Social Psychology, 32, 381-390.

Terman, L.M. (1955) 'Are scientists different?' Scientific American, 192, 25-29.

Tetenbaum, T.J., Lighter, J. and Travis, M. (1981) 'Educators' attitudes toward working mothers' Journal of Educational Psychology, 73, 369-375.

Thomas, W.I. (1928) The child in America, New York, Knopf.

Thompson, N. (1979) 'Sex differentials in physics education' Physics Education, 14, 285-288.

Thorndike, R.L. (1968) Review of 'Pygmalion in the classroom' by R. Rosenthal and L. Jacobson. American Educational Research Journal, 5, 708-711.

Tibbetts, S. (1975) 'Children's literature - a feminist viewpoint' California Journal of Educational Research, 26, 1-5.

Touhey, J.C. (1974) 'Effects of additional women professionals on ratings of occupational prestige and desirability' Journal of Personality and Social Psychology, 29, 86-89.

Travis, C.B. (1982) 'Sex comparisons on causal attributions: Another look at the null hypothesis' Sex Roles, 8, 375-380.

Triandis, H.C. (1971) Attitude and attitude change, New York, John Wiley.

Tuckman, B.W. (1978) Conducting educational research (2nd edition), New York, Harcourt Brace Jovanovich.

Tysoe, M. (1982) 'Do colleges mark women down?' New Society, 9 Dec. pp. 429-430.

United States Office of Education (1976) Women's Educational Equity Act Program Rules and Regulations, Federal Register, Washington, D.C.

Valle, V.A. and Frieze, I.H. (1976) 'The stability of causal attributions as a mediator in changing expectations for success' Journal of Personality and Social Psychology, 33, 579-587.

Vockell, E.L. and Lobonc, S. (1981) 'Sex-role stereotyping by high school females in science' Journal of Research in Science Teaching, 18, 209-219.

Wade, B. (1978) 'Responses to written work: The possibilities of utilizing pupils' perceptions' Educational Review, 30, 149-158.

Walberg, H.J. (1969) 'Physics, femininity and creativity' Developmental Psychology, 1, 47-54.

Walford, G. (1980) 'Sex bias in physics textbooks' School Science Review, 62, No. 219, 220-227.

Walford, G. (1981a) 'Do chemistry textbooks present a sex-biased image?' Education in Chemistry, 18, 18-19.

Walford, G. (1981b) 'Tracking down sexism in physics textbooks' Physics Education, 16, 261-265.

Wallis, W.A. (1942) 'Compounding probabilities from independent significance tests' Econometrica, 10, 229-248.

Wallsgrove, R. (1980a) 'Sisters in science' Science for People, No. 47, p. 16-19.

Wallsgrove, R. (1980b) 'The masculine face of science' in Alice through the microscope by The Brighton Women and Science Group, London, Virago.

Ward, C.A. (1977) Explorations in sex role stereotypes, PhD thesis, University of Durham.

Ward, C. (1979a) 'Differential evaluation of male and female expertise: Prejudice against women?' British Journal of Social and Clinical Psychology, 18, 65-69.

Ward, C. (1979b) 'Explorations in sex role stereotypes' Dissertation Abstracts International, 39, 3598B.

Ward, C. (1981a) 'Prejudice against women: who, when, and why?' Sex Roles, 7, 163-171.

Ward, C. (1981b) 'The differential assessment of males and females as prospective employees by a sample of nonworking women' Sex Roles, 7, 811-820.

Weiner, B. (1974) 'Achievement motivation as conceptualized by an attribution theorist' in B. Weiner (ed) Achievement motivation and attribution theory, Morristown, New Jersey, General Learning Press.

Weiner, B. (1979) 'A theory of motivation for some classroom experiences' Journal of Educational Psychology, 71, 3-25.

Weiner, B., Frieze, I., Kukla, A. Reed, L., Rest, S. and Rosenbaum, R.M. (1971) Perceiving the causes of success and failure, New York, General Learning Press.

Weiner, B., Heckhausen, H., Meyer, W. and Cook, R.E. (1972) 'Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control' Journal of Personality and Social Psychology, 21, 239-248.

Weiner, B., Russell, D. and Lerman, D. (1978) 'Affective consequences of causal ascriptions' in J.H. Harvey, W.J. Ickes and R.F. Kidd (eds) New directions in attribution research, Vol. 2, Hillsdale, New Jersey, Lawrence Erlbaum Associates.

Weinreich-Haste, H. (1979) 'What sex is science?' in O. Hartnett, G. Boden and M.Fuller (eds) Sex-role stereotyping, London, Tavistock.

Weinreich-Haste, H. (1981) 'The image of science' in A. Kelly (ed) The missing half: Girls and science education, Manchester University Press.

Weitzman, L.J. Eifler, D. Hokada, E. and Ross, C. (1972) 'Sex-role socialization in picture books for preschool children' American Journal of Sociology, 77, 1125-1150.

Wells, W.D. and Smith, G. (1960) 'Four semantic rating scales compared' Journal of Applied Psychology, 44, 393-397.

West, C.K. and Anderson, T.H. (1976) 'The question of preponderant causation in teacher expectancy research' Review of Educational Research, 46, 613-630.

Westermann, R. and Hager, W. (1983) 'The relative importance of low significance level and high power in multiple tests of significance' Perceptual and Motor Skills, 56, 407-413.

Whyld, J. (1980) 'Girls and science education' Science for People, 46, 6-8.

Whyld, J. (1983) 'School life - organization and control' in J. Whyld (ed) Sexism in the secondary curriculum, London, Harper and Row.

Whyte, J. (1981) 'Sex typing in schools' in A. Kelly (ed) The missing half: Girls and science education, Manchester University Press.

Whyte, J. (1983a) 'Courses for teachers on sex differences and sex typing' Journal of Education for Teaching, 9, 235-248.

Whyte, J. (1983b) Non-sexist teachers: Evaluating what teachers can do to help girls to opt in to science and technology. Paper presented at the second Girls and Science and Technology Conference, Oslo.

Whyte, J. and Smail, B. (1982) 'GIST - as action research' EOC Research Bulletin, No. 6, Manchester, Equal Opportunities Commission.

Williams, J.A. (1977) Psychology of women, New York, Norton.

Williams, J.E., Giles, H., Edwards, J.R., Best, D.L. and Daws, J.T. (1977) 'Sex-trait stereotypes in England, Ireland and the United States' British Journal of Social and Clinical Psychology, 16, 303-309.

Williams, T. (1976) 'Teacher prophecies and the inheritance of inequality' Sociology of Education, 49, 223-236.

Wilson, G.D. (ed) (1973) The Psychology of Conservatism, London, Academic Press.

Wilson, G.D. and Patterson, J.R. (1968) 'A new measure of conservatism' British Journal of Social and Clinical Psychology, 7, 264-269.

Wilson, W.L. (1974) 'The effect of teacher-pupil sex interaction and teacher age on pupil perception of teacher behavior in junior and senior high school classes of English in selected Mississippi public schools' Dissertation Abstracts International, 34, 5004A.

Winkler, A. (1982) 'The effects of sex and interpersonal dominance on evaluations of nurses' competence and attractiveness' International Journal of Women's Studies, 5, 97-113.

Wiseman, S. and Wrigley, J. (1958) 'Essay-reliability: The effect of choice of essay-title' Educational and Psychological Measurement, 18, 129-138.

Wittekind, O.S.F. (1975) 'Relationships between self-esteem, sex-type, and perceived competency of 'male' and 'female' authors' Dissertation Abstracts International, 36, 6-B, 3081.

Wolpe, A. (1977) Some processes in sexist education, London, WRRC.

Youngman, M.B. (1978) Designing and analysing questionnaires, Rediguide 12, Nottingham University School of Education.

Youngman, M.B. (1979) *Presenting research results*, Rediguide 25, Nottingham University School of Education.

Youngman, M.B. 'Correlation scatter plotting as a quick substitute for factor analysis' *Research in Education*, No. 25, 13-18.

Zimet, S.G. and Zimet, C.N. (1977) 'Teachers view people: Sex role stereotyping' *Psychological Reports*, *41*, 583-591.